Cancellation, Negation, and Rejection

Niels Skovgaard-Olsen

University of Göttingen

Peter Collins

Goldsmiths, University of London

Karolina Krzyżanowska

University of Amsterdam

Ulrike Hahn

Birkbeck, University of London

Karl Christoph Klauer

Albert-Ludwigs-Universität Freiburg

Author Note

Niels Skovgaard-Olsen, Department of Cognition and Decision Making, University of Göttingen, Germany. Peter Collins, Department of Psychology, Goldsmiths, University of London. Karolina Krzyżanowska, Institute for Logic, Language and Computation, University of Amsterdam, The Netherlands. Ulrike Hahn, Department of Psychological Sciences, Birkbeck, University of London, UK. Karl Christoph Klauer, Department of Psychology, Albert-Ludwigs-Universität Freiburg, Freiburg, Germany.

This work was supported by grants to Wolfgang Spohn and Karl Christoph Klauer from the Deutsche Forschungsgemeinschaft (DFG) as part of the priority program "New Frameworks of Rationality" (SPP 1516).

Correspondence concerning this article should be addressed to Niels Skovgaard-Olsen (niels.skovgaard-olsen@psych.uni-goettingen.de, n.s.olsen@gmail.com).

Supplementary Materials: https://osf.io/hz4k6/

Abstract

In this paper, new evidence is presented for the assumption that the reason-relation reading of indicative conditionals ('if A, then C') reflects a conventional implicature. In four experiments, it is investigated whether relevance effects found for the probability assessment of indicative conditionals (Skovgaard-Olsen, Singmann, and Klauer, 2016a) can be classified as being produced by a) a conversational implicature, b) a (probabilistic) presupposition failure, or c) a conventional implicature. After considering several alternative hypotheses and the accumulating evidence from other studies as well, we conclude that the evidence is most consistent with the Relevance Effect being the outcome of a conventional implicature. This finding indicates that the reason-relation reading is part of the semantic content of indicative conditionals, albeit not part of their primary truth-conditional content.

Keywords: Relevance, Pragmatics, Indicative Conditionals, Conversational Implicatures, Presuppositions, Conventional Implicatures

Cancellation, Negation, and Rejection[1]

Very few linguistic expressions are as important to our reasoning, argumentation, and decision making as indicative conditionals, that is, a class of sentences typically of the form: "If A, (then) C," where A, the *antecedent*, and C, the *consequent*, stand for arbitrary sentences. It is not surprising then that conditionals have been a subject of extensive research in philosophy, linguistics, computer science, and psychology. What is more surprising though is that, despite decades of multidisciplinary efforts to understand how people interpret indicative conditionals, many aspects of their meaning remain the matter of ongoing controversy.

Intuitively, whenever someone asserts a conditional they communicate that there is some sort of a relation between the content of that conditional's antecedent and consequent; that, for instance, the antecedent is a reason for the consequent, or that the consequent can be inferred from the antecedent. Take, for example, the following sentence:

(1) If more parents refuse to vaccinate their children, diseases such as measles and whooping cough will make a comeback.

Clearly, someone who asserts (1) seems to be expressing their belief about the connection between the anti-vaccination movement and the possible outbreak of infectious diseases. That a conditional conveys such a relationship is not controversial. But the status of this connection is one of the most contentious issues in the current debate on the meaning of indicative

[1] Acknowledgement: We would like to thank our Editor, Art Markman, and reviewers for detailed comments that helped improve the paper. In particular, David Over, Florian Schwarz, and Sunny Khemlani deserve credit. Furthermore, we would like to thank audiences to talks at New Perspectives on Conditionals and Reasoning (Regensburg, 2018), Annual Meeting of New Frameworks of Rationality (Etelsen, 2018), What if, Scientific Retreat (Konstanz, 2018), ECAP9: European Congress of Analytic Philosophy (Munich, 2017), as well as Robyn Carston, John MarFarlane, Maria Biezma, and others, who were so helpful as to engage in discussions with us over these matters. Thanks also go to students helping us carrying out the experiments like Dennis Heiler.

conditionals. At issue is whether the connection is part of the *semantics* or the *pragmatics* of the conditional. Roughly, the semantics of the conditional – sometimes referred to as the *core meaning* – is its literal, conventional, context-independent meaning. The pragmatics of the conditional is its non-literal, inferred, context-dependent meaning (we will elaborate on these definitions later).

The question of whether the connection is semantic or pragmatic has attracted such interest in large part because the issue forms a dividing line between theories of the conditional in psychology and philosophy. On the one hand, there are historically established theories in psychology, such as the Suppositional Theory and the Mental Models Theory, which take the connection to be pragmatic. On the other, there are recently revived 'inferentialist' accounts, which take the connection to be semantic. The opposition of these theories in psychology is an echo of similar debates in philosophy (for references, see Douven, 2015 and Skovgaard-Olsen, 2016).

In comparison, linguistic debates about conditionals have been more influenced by the possible-worlds semantics of Stalnaker (1968), Lewis (1973), and Kratzer (1986). In none of these theories is the connection between the antecedent and the consequent taken to be semantic. Rather what we get is roughly a description of the antecedent worlds in the context set, which are most similar to the actual world, stating that they are worlds in which the consequent is true (Biezma, 2014). As such, formal semantics sides with the psychological theories in denying that the connection between the antecedent and consequent is part of the core semantic meaning of conditionals. But in this paper, our focus is on the former division within theories of conditionals in psychology. See, however, Skovgaard-Olsen (*in review*) for more on the connections between the present discussion and related work in linguistics.

The Mental Model Theory is one of the most influential theories in the psychology of reasoning. It postulates that compound sentences, such as conditionals, refer to conjunctions of possibilities, where possibilities are understood epistemically as situations that are compatible with what is known (Khemlani, Byrne, and Johnson-Laird, 2018). Consequently, on the Mental Model Theory, interpreting a sentence amounts to constructing mental models that represent possible states of affairs that are compatible with that sentence, while what is impossible tends to be omitted (Johnson-Laird and Byrne 1991, 2002; Johnson-Laird, Khemlani, & Goodwin 2015). A fully fleshed out, explicit model of a conditional, if A then C, can be depicted then in the following way, where each row denotes a mental model of one possibility:[2]

A C

¬A ¬C

¬A C

Importantly, already Johnson-Laird and Byrne (1991, 2002), but also Khemlani et al. (2018), emphasized that many people do not immediately construct fully explicit models of a conditional. Instead they stop their model construction at the initial, abbreviated, *implicit* model consisting of the representation of the possibility that both the antecedent and the consequent are true:

A C

. . .

The ellipsis signals that there are other possibilities, which could be evoked, if necessary.

[2] For consistency we use 'A' and 'C' throughout to refer to the antecedent and consequent of the conditional respectively.

It is important to note that such conjunctions of possibilities constitute the meaning of a compound sentence on the Mental Models Theory. More precisely, the Mental Model Theory holds that a sentence is true if all the corresponding models are possible. This makes the mental models of a conditional different than a material conditional, even though the explicit model bears a resemblance to the truth table for the material implication. As Khemlani et al. (2018) say:

> In the model theory, a conditional's meaning is not a material implication, not a conditional probability, not a set of possible worlds, and not an inferential relation. It is instead a conjunction of possibilities, each of which is assumed in default of information to the contrary. (p. 31)

Accordingly, Johnson-Laird et al. (2015) hold that "a basic conditional, 'if A then C', is true only if all three situations in its fully explicit models are possible: "possibly(A & C) & possibly(not-A & not-C) & possibly(not-A & C) and A & not-C is impossible" (p. 206).[3]

But what matters for present purposes is just that although Mental Model Theory does not treat the relation between the antecedent and consequent of a conditional as a part of its core meaning, proponents of Mental Model Theory acknowledge that conditionals are often interpreted as conveying that there is such a relation due to "modulating effects of semantics and pragmatics" (Johnson-Laird and Byrne, 2002, p. 651). More specifically, "the meaning of words, knowledge, and the conversational context can block the construction of models of possibilities, and they can add causal, spatiotemporal, and other relations between elements in models (Khemlani et al., 2018, p. 12-13). They argue that the context of an utterance (pragmatics), or

[3] See Baratgin, Douven, Evans, Oaksford, and Politzer (2015) for a discussion of some challenges for the revised version of the theory, and Khemlani et al. (2018) for a response. See further the discussion in Hinterecker, Knauff, and Johnson-Laird (2016) and Oaksford, Over, and Cruz (2018).

semantic relationship between the content of the antecedent and the content of the consequent, may block the construction of a model that normally belongs to the core meaning of a sentence, or trigger the construction of a model that is not part of the core meaning of a sentence. As we will argue below, semantic and pragmatic modulation is not sufficient to account for the complex data pattern that emerges out of the present study when taken together with other recent published results, however.

An alternative approach in the psychological study of conditionals stems from the so-called New Paradigm Psychology of Reasoning (Oaksford & Chater, 2007; Over, 2007), which emphasizes the role of uncertainty in human reasoning. Indicative conditionals on the New Paradigm are interpreted as probabilistic or suppositional. Although the term *Suppositional Theory* (ST) refers to a whole family of related views, they can all be construed as the formalizations of the Ramsey Test, which provides a procedure for fixing one's degree of belief in a conditional, and, by the same token, for determining whether an indicative conditional is acceptable (Ramsey 1929/1990, p. 155):

> If two people are arguing 'if [A] will [C]?' and are both in doubt as to [A], they are adding [A] hypothetically to their stock of knowledge and arguing on that basis about [C]: so that in a sense 'If [A], [C]' and 'if [A], [¬C]' are contradictories. We can say that they are fixing their degrees of belief in [C] given [A]. (Editorial changes preserve the consistency of notation)

Consequently, what the different versions of the Suppositional Theory have in common is their commitment to *The Equation,* according to which the probability of a conditional equals the conditional probability of that conditional's consequent given its antecedent (where 'A' and 'C' are restricted to atomic sentences, that is, they are not conditionals themselves): P(If A, then C) =

P(C|A). The development of the Suppositional Theory that became particularly influential in the psychology of reasoning resulted from combining the Ramsey Test, and thus the Equation, with three-valued de Finetti's semantics. De Finetti treated conditionals as true when both of its clauses are true, and false when the antecedent is true but the consequent is false. When the antecedent is false, the truth value of a conditional is undetermined, "void." Conditionals with false antecedents can be compared to called-off bets: a bet that if you throw a fair coin it will land heads is neither won, nor lost – it is called off – when the coin is not thrown at all (see, e.g., Politzer, Over, & Baratgin, 2010). A more refined version of the de Finetti's system can be obtained by replacing the third, "void" value with the conditional probability itself (Jeffrey, 1991; see also Baratgin, Politzer, Over, and Takahashi, 2018; Kleiter, Fugard, & Pfeifer, 2018, and Over & Cruz, 2018).

The Equation has received strong empirical support. Participants in reasoning experiments tend to evaluate the probability of a conditional by estimating the corresponding conditional probability (Evans, Handley, & Over, 2003; Fugard, Pfeifer, Mayerhofer, & Kleiter, 2011; Oberauer & Wilhelm, 2003; Politzer et al., 2010). Yet, a recent study has challenged the generality of these results (Skovgaard-Olsen, Singmann, and Klauer, 2016a), as we shall see in more detail below.

Although the original phrasing of the Ramsey Test, with its focus on arguing about C *on the basis of* A, seems to capture the intuition that the antecedent of a conditional needs to be somehow relevant for the consequent, this is not true of the suppositional accounts. As long as the antecedent is possible, a true or even highly probable consequent will render the conditional acceptable. If we believe that Brexit is inevitable and that it is quite possible that there are at

least some microorganisms living on some planets outside of our Solar System, we are committed to accepting the following missing-link conditional:

(2) If there is life on some extra-solar planet, then the UK will leave the European Union.

This is because "the UK will leave the European Union" was already part of our stock of beliefs, and it remains so upon expanding it by "there is life on some extra-solar planet." By contrast, for a person who deems it rather unlikely that there are any advanced alien civilizations, and that this likelihood will not increase were we to learn that there are planets that host some form of life, (3) does not appear acceptable after performing the Ramsey Test:

(3) If there is life on some extra-solar planet, then somewhere in the Universe there exists an advanced alien civilization.

Yet it would be easier to make sense of a speaker who asserts (3), even if we strongly disagree with it, than of someone who asserts (2) (Krzyżanowska, 2015, p. 9). If we disagree with (3), it would be because learning that there is life on some extra-solar planet would not be a good enough reason for us to believe in the existence of an advanced alien civilization. Nevertheless, since the truth of the antecedent of (3) slightly increases the probability of its consequent, we can imagine someone who would find such an argument convincing enough to accept the conditional.

Psychologists of reasoning who claim that the Equation captures a central part of the meaning of an indicative conditional do not deny that these sentences often seem to suggest stronger (e.g. causal or inferential) relations between their antecedents and consequents. What they do deny is that these relations belong to the semantics, that is, to the core meaning of the conditional. It seems that when researchers wish to account for our intuitions about the relation between a conditional's antecedent and consequent, and for data on the actual use of conditionals,

they most commonly invoke pragmatics (Johnson-Laird & Byrne, 2002; Over et al., 2007; for discussion, see, also Krzyżanowska, Collins, and Hahn, 2017a, 2017b, and below). But pragmatics is not the only option. The alternative approach is to treat that relation, however it is defined, as the starting point for developing an account of the meaning of a conditional. Douven and Verbrugge (2010) argued that one can distinguish between different types of inferences, and classified inferential conditionals as *deductive*, *inductive*, or *abductive*, inspired by classifications of conditionals in empirical linguistics (Declerck and Reed, 2001; Dancygier and Sweetser, 2005). This typology motivated a new, "inferential," truth-conditional semantics for indicative conditionals (Krzyżanowska, Wenmackers, and Douven, 2014; Douven, 2015). Independently, a related account has been proposed by Spohn (2013) and Olsen (2014), who analyzed indicative conditionals as expressing *reason relations* between their antecedents and consequents. This relationship can be operationalized probabilistically in terms of the ΔP rule ($\Delta P = P(C|A) - P(C|\neg A)$). A conditional's antecedent, A, is said to be a reason for the consequent, C, if A raises the probability of C, that is, if A is positively relevant for C. Since ΔP is defined as a difference between $P(C|A)$ and $P(C|\neg A)$, ΔP must be positive for A to be a reason for C, and, consequently, for a conditional, "If A, then C," to be acceptable. Positive Relevance can be seen in example (1) above: parents' refusing to vaccinate their children increases the probability of measles or whooping cough outbreaks. By contrast, probabilistic irrelevance can be seen in example (2) above: the probability of UK leaving the EU given the existence of life on some extra-solar planet is, to the best of our current knowledge, exactly the same as the probability of UK leaving the EU given that there is no life outside of the Solar system at all. That is, $\Delta P = 0$, or the antecedent is probabilistically irrelevant for the consequent in this case.

The Relevance Effect

Results by Skovgaard-Olsen et al. (2016a) recently raised an explanatory challenge for the Suppositional Theory of conditionals and Mental Model Theory. Both theories postulate that indicative conditionals have a core meaning which does not include relevance relations between the antecedent and the consequent. However, when investigating the probability and acceptability of indicative conditionals, Skovgaard-Olsen et al. (2016) found that relevance strongly moderated the evaluations of indicative conditionals. For cases of Positive Relevance ($P(C|A) - P(C|\neg A) > 0 \Leftrightarrow \Delta P > 0$), the conditional probability remained a good predictor of both the acceptance and probability of the conditional. For cases of Negative Relevance ($P(C|A) - P(C|\neg A) < 0 \Leftrightarrow \Delta P < 0$) and Irrelevance ($P(C|A) - P(C|\neg A) = 0 \Leftrightarrow \Delta P = 0$), this relationship was disrupted because the participants tended to view the indicative conditional as defective under those conditions.

In what sense does the Relevance Effect constitute a challenge to the Mental Model theory and the Suppositional Theory? The extent to which it does depends on whether the Relevance Effect belongs to the core meaning of the conditional - its semantics - or arises, instead, from the context of utterance of a conditional - its pragmatics. If the Relevance Effect belongs to pragmatics, then the main theories can just claim to be theories about the core content of indicative conditionals and hold that they need to be supplemented with auxiliary hypotheses concerning the pragmatic mechanisms involved in communication.

To address this question we focus on a set of well-known phenomena at the interface between semantics and pragmatics: namely conversational implicature, presupposition, and conventional implicature (see later sections for definitions of each). We do so because these are phenomena for which there are reasonably well-established diagnostic tests. If we can explain

the Relevance Effect by one of these phenomena, we are a step closer to adjudicating on the semantics/pragmatics issue. A final judgment will depend both on how we define semantics and pragmatics and on how we subsequently classify conversational implicature, presupposition, and conventional implicature. Both the definition and subsequent classification are live issues. But instead of resolving those issues here, our focus will be on classifying relevance effects with respect to these three established linguistic phenomena at the interface between semantics/pragmatics. For present purposes, we follow Birner (2014) in adopting the following typical characteristics of semantics and pragmatics (Table 1):

Table 1. Pragmatic/Semantic Distinction

Semantics	*Pragmatics*
literal	non-literal
context-independent	context-dependent
non-inferential	inferential
truth-conditional	non-truth-conditional[4]

To this we might add that semantics typically concerns the conventional meaning of words and sentences, while pragmatics typically concerns non-conventional meaning. While these characteristics might define the prototypical semantic and pragmatic phenomena, the characteristics can come apart. For instance, a phrase such as 'the foot of the mountain' may strike us as non-literal, mountains not having body parts, but it will also likely strike us as non-inferential, truth-conditional, and conventional. Unsurprisingly, then, it can prove controversial to categorize any given phenomenon as semantic or pragmatic. Of the phenomena we consider, only conversational implicatures are regarded as uncontroversially pragmatic. Conventional implicatures, in contrast, are commonly thought of as a secondary layer of semantic meaning

[4] Some would argue that some pragmatic phenomena are, in fact, truth-conditional. For discussion, see Carston (2002), Recanati (2011), and Birner (2014).

which is auxiliary to the primary truth-conditional semantic layer (Potts, 2007, 2015). Presuppositions, on the other hand, have both semantic and pragmatic interpretations (Beaver & Geurts, 2014), with influential proponents on either side—with, for instance, Stalnaker (2016) defending a pragmatic approach and von Fintel (2008) adopting a semantic one. The distinction between these various linguistic categories is discussed in further details below.

We test among the linguistic categories in four experiments. Experiment 1 tests whether the Relevance Effect arises because of conversational implicature. Experiment 2 tests whether it arises because of a presupposition failure. Experiment 3 tests whether it arises because of a conventional implicature.

Experiment 1

Conversational Implicatures

We start with the paradigm-case of pragmatics: the conversational implicature. Conversational implicatures arise when a speaker means something different from the conventional meaning of the sentence they utter. For instance:

Alan: Are you going to Paul's party?

Barb: I have to work. (Davis, 2014)

Here, Barb utters a sentence with a clear conventional meaning—that she has to work—but also conversationally implicates that she will not be attending Paul's party (because she has to work). To take another familiar example:

Angry Parent: Did you eat all of the chocolate cake?

Guilty Child: I ate *some* of it.

Here, the child utters a sentence that conventionally means something like 'I ate at least one morsel of cake' and is quite compatible with 'I ate all of the cake'. But the child, perhaps hoping

to spread the blame, also implicates that he/she did not eat all of the cake: that there is another culprit. Inferences of this latter type are known as scalar implicatures.

Grice (1989) set out to explain how conversational implicatures arise, formulating a general principle of cooperative discourse: that speakers 'make [their] contribution such as is required, at the stage at which it occurs, by the accepted purpose or direction of the talk exchange in which [they] are engaged' (Grice, 1975, p. 45). He fleshed this principle out into a set of conversational maxims, or descriptive norms. On his account, speakers should give enough, but not too much, information (Maxim of Quantity); should avoid saying falsehoods or things for which they lack evidence (Maxim of Quality); should be relevant (Maxim of Relation); and should avoid obscurity and ambiguity, and be brief and orderly (Maxim of Manner).

According to Grice, implicatures can only arise at all because hearers assume that speakers are generally cooperative: that they follow the maxims. But, as he pointed out, speakers can, in fact, behave in different ways towards the maxims: they can observe, violate, flout, or opt out of a maxim. Most important for present purposes are the observing or flouting of maxims, either of which generates an implicature.[5] For instance, other things being equal, when a speaker says 'I had two bagels for breakfast', a hearer will assume that the speaker is observing the Maxim of Quantity (is providing sufficient information) and implicating 'I did not have three bagels for breakfast' (Birner, 2013). For an example of flouting, consider a professor who is writing a recommendation letter for a student, and is expected, in the normal run of things, to

[5] A speaker violates a maxim when they inconspicuously disregard it, not intending the hearer to notice—as when a speaker lies or misleads a hearer; a speaker opts out of a maxim when, say, they simply disengage from a conversation and 'do not play the game at all' (Birner, 2013).

comment favorably on the student's academic ability, diligence, and so on. A professor who comments, instead, on the student's handwriting is flouting—openly disregarding—the Maxim of Quantity and implicating that the student's academic ability (and so on) is not worthy of praise (Grice, 1989). In cases of flouting, we can think of an implicature as being necessary to preserve the assumption that the speaker is being cooperative.

Conversational implications constitute a paradigmatic example of pragmatic modulation, an interpretational process whereby semantic content is extended by pragmatic mechanisms that take contextual factors into account. It is not unusual to find references to processes of pragmatic modulation as extending semantic theories in the psychology of reasoning (e.g. Johnson-Laird and Byrne, 2002). But their status tends to be that of underspecified and untested, auxiliary hypotheses (Douven, 2017).

Since conversational implicatures are attempts to reconstruct the speaker's intended meaning, which goes beyond what is literally said, they are defeasible inferences, which can be explicitly blocked by the speaker. For instance, imagine a conversation between John and Sophia at a party they are hosting. John, who is in the kitchen, asks Sophia 'Where are our guests?' If Sophia replies 'Some are in the garden' she might well be taken to mean that not all of the guests are in the garden. But Sophia can straightforwardly cancel this scalar implicature by adding, 'In fact, they all are.' We will call this type of hedging 'a cancellation speech act'. One of the main characteristics of conversational implicatures is that a commitment to them can be blocked by performing cancellation speech acts without producing a contradiction (Blome-Tillmann, 2013).

Conversational implicatures have already featured in the debate on relevance. For instance, Over et al. (2007) found a modest effect of $P(C|\neg A)$ as a predictor of P(if A, then C),

for conditionals which were positively relevant (where $\Delta P > 0$). They offered the following explanation (p. 92):

> An Adams conditional is not equivalent to an explicit statement that [A] raises the probability of [C] (...), nor that [A] causes [C] (...). A conditional probability [P(C|A)] can be high when [A] does not raise the probability of [C] and when [A] does not cause [C]. For example, [P(C|A)] can be high simply because [P(C)] is high. Does this mean that supporters of the view that these conditionals are Adams conditionals cannot account for the weak negative effect of [P(C| ¬A)] in the current studies? Not necessarily, for they can argue that the use of a conditional pragmatically suggests, in certain ordinary contexts, that [A] raises the probability of [C] or that [A] causes [C]. (Editorial changes preserve the consistency of notation),

Over et al. (2007) then go on to suggest that probability raising and the causal reading of indicative conditionals may be produced by a conversational implicature. It may be misleading to assert conditional sentences in the absence of probabilistic dependencies, but the reason for this does not reside in the core, semantic content of indicative conditionals.

Similarly, Johnson-Laird and Byrne (2002) have considered whether to make a connection between the antecedent and the consequent part of the core meaning of conditionals, only to reject it:

> We do not deny that many conditionals are interpreted as conveying a relation between their antecedents and consequents. However, the core meaning alone does not signify any such relation. If it did, then to deny the relation while asserting the conditional would be to contradict oneself. Yet, the next example is not a contradiction:

> If there was a circle on the board, then there was a triangle on the board, though there was no relation, connection, or constraint, between the two—they merely happened to co-occur. (p. 651)

Their argument is that one can cancel a commitment to there being a relation between the antecedent and the consequent without contradicting oneself. If so, then this commitment bears the mark of a conversational implicature. In Experiment 1 we will test this hypothesis.

The Cancellation Task

The purpose of Experiment 1 is to test whether the reason-relation reading of conditionals can be attributed to the presence of a conversational implicature. To test this hypothesis, we investigated whether the reason-relation reading of conditionals can be cancelled without contradiction. More specifically, Experiment 1 uses the perceived degree of contradiction in cancelling a scalar implicature as the lower baseline. We have already seen two examples of scalar implicatures, where speakers used (or could be mistaken for using) the weaker term 'some' to implicate 'not all'. The 'some' to 'not all' inference is the most famous case, but scalar implicatures can arise with various scales, such as scales of possibility ('It's possible he will come' can implicate 'It's not definite that he'll come'). The implicature is that the speaker has some reason for not using the more informative, stronger term. Scalar implicatures can be cancelled. In this respect they contrast markedly with our upper baseline, entailment. If sentence A entails sentence B, then whenever A is true, B is also true. For instance, 'John is a bachelor' entails 'John is unmarried'. By definition, entailments cannot be cancelled without contradiction. The test then consists in measuring whether attempts to cancel the reason-relation reading of conditionals are viewed as more like cancelling a scalar implicature than like cancelling an

entailment relation. The rationale is that while scalar implicatures can be cancelled without contradiction, entailments cannot.

In addition, Experiment 1 contrasts attempts to cancel the reason-relation reading of conditionals with attempts to cancel the reason-relation reading of conjunctions, which is another connective featuring a prominent reason-relation reading (Carston, 1993). Finally, comparisons are made with two control items that do not involve conditionals. The first is an attempt to cancel a scalar implicature. The second is an attempt to cancel the entailment of a categorical assertion.

Method

Participants

The experiment was conducted over the Internet to obtain a large and demographically diverse sample. A total of 100 people completed the experiment. The participants were sampled through the Internet platform Mechanical Turk from the USA, UK, Canada, and Australia. They were paid a small amount of money for their participation.

The following exclusion criteria were used: not having English as native language (zero participants), completing the task in less than 240 seconds or in more than 3600 seconds (15 participants), failing to answer two simple SAT comprehension questions correctly in a warm-up phase (44 participants), and answering ‘not serious at all’ to the question 'how serious do you take your participation' at the beginning of the study (1 participant). Since some of these exclusion criteria were overlapping, the final sample consisted of 65 participants. Mean age was 41.66 years, ranging from 24 to 65, 38.5 % of the participants were male; 58.5 % indicated that the highest level of education that they had completed was an undergraduate degree or higher. Applying the exclusion criteria had a minimal effect on the demographic variables.

Design

The experiment had a within-subject design with three factors: Relevance (with two levels: Positive Relevance (PO), Irrelevance (IR)), Priors (with four levels: HH, HL, LH, LL, meaning, for example, that P(A) = low and P(C) = high for LH), and Sentence Type (with two levels: Conditional (if), Conjunction (and)). Since the And_{PO} cell of our design was empty,[6] the participants were presented with 12 within-subject conditions.

Materials and Procedure for All the Experiments

Each of the 12 within-subjects conditions was randomly assigned one of 12 scenarios. Random assignment was performed without replacement such that each participant saw a different scenario for each condition. This ensured that the mapping of condition to scenario was counterbalanced across participants preventing confounds of condition and content. The list of the 12 scenarios used can be found in the Supplemental Materials.

To reduce the dropout rate during the experiment, participants first went through three pages stating our academic affiliations, posing two SAT comprehension questions in a warm-up phase, and presenting a seriousness check asking how careful the participants would be in their responses (Reips, 2002).

The experiment was split into twelve blocks of four pages, one block for each within-subjects condition. The order of the blocks was randomized anew for each participant and there were no breaks between the blocks. On the first page of each block, the participants were presented with a randomly chosen scenario text (which was repeated on the three following pages in a brighter grey color for reference). These scenario texts have been found in previous studies (Skovgaard-Olsen, Singmann, and Klauer, 2016b) to reliably induce assumptions about

[6] To avoid prolonging the experiment too much for the participants, we chose to focus on the IF_{IR} and AND_{IR} comparison in this paper.

relevance and prior probabilities of the antecedent and the consequent that implement our experimental conditions. Table 2 displays sample conditions for the Mark scenario for Positive Relevance and Irrelevance.

Table 2. Stimulus Materials, Mark Scenario

Scenario	Mark has just arrived home from work and there will shortly be a great movie on television, which he has been looking forward to. Mark is quite excited because he recently bought a new TV with a large screen. He has a longing for popcorn, but his wife has probably eaten the last they had while he was gone.			
	Positive Relevance		**Irrelevance**	
HH	If Mark presses the on switch on his TV, then his TV will be turned on.		If Mark is wearing socks, then his TV will work.	
HL	If Mark looks for popcorn, then he will be having popcorn.		If Mark is wearing socks, then his TV will malfunction.	
LH	If the sales clerk in the local supermarket presses the on switch on Mark's TV, then his TV will be turned on.		If Mark is wearing a dress, then his TV will work.	
LL	If Mark pulls the plug on his TV, then his TV will be turned off.		If Mark is wearing a dress, then his TV will malfunction.	
	Positive Relevance (PO):	mean ΔP = .32	High antecedent:	mean P(A) = .70
	Irrelevance (IR)	mean ΔP = -.01	Low antecedent:	mean P(A) = .15
			High consequent:	mean P(C) = .77
			Low consequent:	mean P(C) = .27

Note. HL: P(A) = High, P(C) = low; LH: P(A) = low, P(C) = high. The bottom rows display the mean values for all 12 scenarios pretested in Skovgaard-Olsen et al. (2016b).

For the Mark scenario text in Table 2, participants assume that Mark pressing the on switch raises the probability of his TV turning on, and that both of these sentences have a high prior probability given the scenario (Positive Relevance, HH). Conversely, the participants tend to assume that Mark's wearing socks is irrelevant for whether the TV will work, and that both have a high prior probability (Irrelevance, HH).

Materials and Procedure specific to Experiment 1

For Experiment 1, the participants were given the following instruction:

> In the following you are going to see a short conversation, where Louis accuses Samuel of contradicting himself. Whether you agree with Samuel's assertions is beside the point. What we are interested in is just the extent to which you agree with Louis that Samuel is contradicting himself. When you read the sentences please pay attention to small differences in their content, so that we don't unfairly accuse Samuel of contradicting himself.

The participants were then presented with two control items:

> Samuel: John is a bachelor [/Some of the employees are invited to the party]
>
> ...but I am not suggesting that John is unmarried. [/that they're not all invited]
>
> Louis: Wait, you're contradicting yourself.

The task of the participants was to indicate the extent to which they agreed or disagreed with Louis's statement on a five-point Likert scale {strongly disagree, disagree, neutral, agree, strongly agree}. Before beginning the experiment proper, the participants moreover saw one practice trial, where we emphasized that attention was needed to notice the subtle differences between the wordings of the various types of cancellations used in the experiment.

On the following three pages, the participants were presented with one of the three dependent variables in random order (perceived contradiction of cancellation of entailment, of scalar implicature, and of the reason-relation reading). The task of the participants was always to assess the extent to which they agreed with Louis' claim that Samuel contradicted himself. Using the HH conditions from above, the three types of cancellation were implemented as follows:

Table 3. Cancellation Types in Experiment 1

Entailment	Scalar Implicature	Reason Relation
	Conditionals, Positive Relevance	
Samuel: Mark presses the on switch on his TV. And IF Mark presses the on switch on his TV, THEN his TV will be turned on. ...but I am not suggesting that Mark's TV will be turned on.	**Samuel:** Mark presses the on switch on his TV. And IF Mark presses the on switch on his TV, THEN it is possible that his TV will be turned on. ...but I am not suggesting that if so, it isn't highly likely that Mark's TV will be turned on.	**Samuel:** Mark presses the on switch on his TV. And IF Mark presses the on switch on his TV, THEN his TV will be turned on. ...but I am not suggesting that these two things are related.
	Conditionals, Irrelevance	
Samuel: Mark is wearing socks. And IF Mark is wearing socks, THEN his TV will work. ...but I am not suggesting that Mark's TV will work.	**Samuel:** Mark is wearing socks. And IF Mark is wearing socks, THEN it is possible that his TV will work. ...but I am not suggesting that if so, it isn't highly likely that Mark's TV will work.	**Samuel:** Mark is wearing socks. And IF Mark is wearing socks, THEN his TV will work. ...but I am not suggesting that these two things are related.
	Conjunctions, Irrelevance	
Samuel: Mark is wearing socks AND his TV will work. ...but I am not suggesting that Mark's TV will work.	**Samuel:** Mark is wearing socks AND it is possible that his TV will work. ...but I am not suggesting that it isn't highly likely that Mark's TV will work.	**Samuel:** Mark is wearing socks AND his TV will work. ...but I am not suggesting that these two things are related.

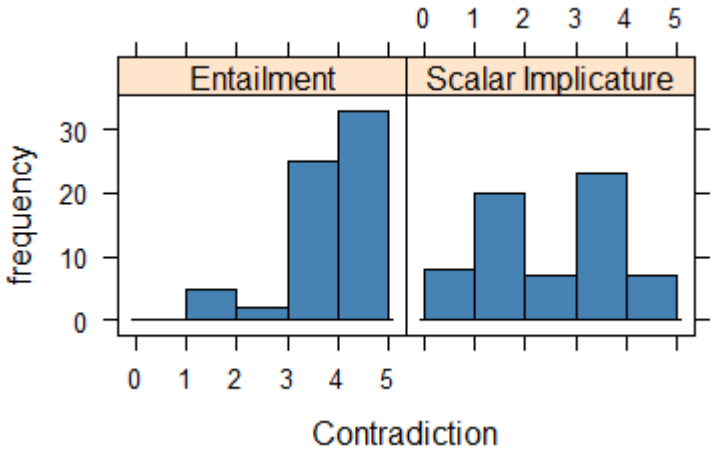
Note. For conditionals, the entailment of the conclusion of Modus Ponens was cancelled. For conjunctions, the entailment of the conclusion of conjunction elimination was cancelled.

Results

Control Items

As Figure 1 suggests, the degree to which the participants viewed the cancellation speech act as giving rise to a contradiction was found to be significantly higher in the entailment control (*Mdn* = 5.00) item than in the scalar control item (*Mdn* = 3.00), V = 1260.5, $p < .0001$, $r = -.65$, for the Asymptotic Wilcoxon signed-rank test.

Figure 1. Histogram for Control Items



Note. The width of the bins is 1, so the bin from 0-1 on the histogram = 'strongly disagree' (or = '1', on the original response scale), the bin from 4-5 on the histogram = 'strongly agree' (or = '5', on the original response scale).

Comparing Cancellation Types for $And_{Irrelevance}$ and $IF_{Positive}$, $IF_{Irrelevance}$

Given the design, there were replicates for each participant and item. Hence, it was not appropriate to assume that the data were independently and identically distributed. Accordingly, the appropriate analysis was to use linear mixed-effects models, with crossed random effects for intercepts and slopes by participants and by scenarios (Baayen, Davidson, and Bates, 2008). This analysis was conducted using the statistical programming language R (R Core Team, 2013), and the package `brms` for mixed-effects models in Bayesian statistics was used (Bürkner, 2017). On the project page on the Open Science Framework, previous analyses of all the experiments reported in this paper are reported for classical statistics: https://osf.io/hz4k6/. As seen, the classical and Bayesian analyses converge on qualitatively similar results for all the studies.

Separate analyses were run for the Irrelevance (IR) and Positive Relevance (PO) items because there was no $AND_{Positive}$ cell of the design. For the Irrelevance items ($AND_{Irrelevance}$, $IF_{irrelevance}$), the following models were fit to the data:

(M1) a model that treats the participants' ratings of perceived contradiction as a function of the Cancellation Type factor (scalar implicature vs. entailment vs. reason relation), the Sentence factor ('if, then' vs. 'and'), and their interaction.

(M2) a model that builds on M1 but without the two-ways interaction.

(M3) a model that builds on M2 but without a main effect for the Sentence factor.

As indicated, these models were implemented in a Bayesian framework with weakly informative priors, using the R package `brms` (Bürkner, 2017). Since the responses obtained from the five-point Likert scale are ordinal responses, the responses were modelled as generated by thresholds set on a latent continuous scale with a cumulative likelihood function and a logit link function (Bürkner & Vuorre, 2018). Table 4 reports the performance of the models as quantified by the leave-one-out cross validation criterion and the WAIC information criterion.

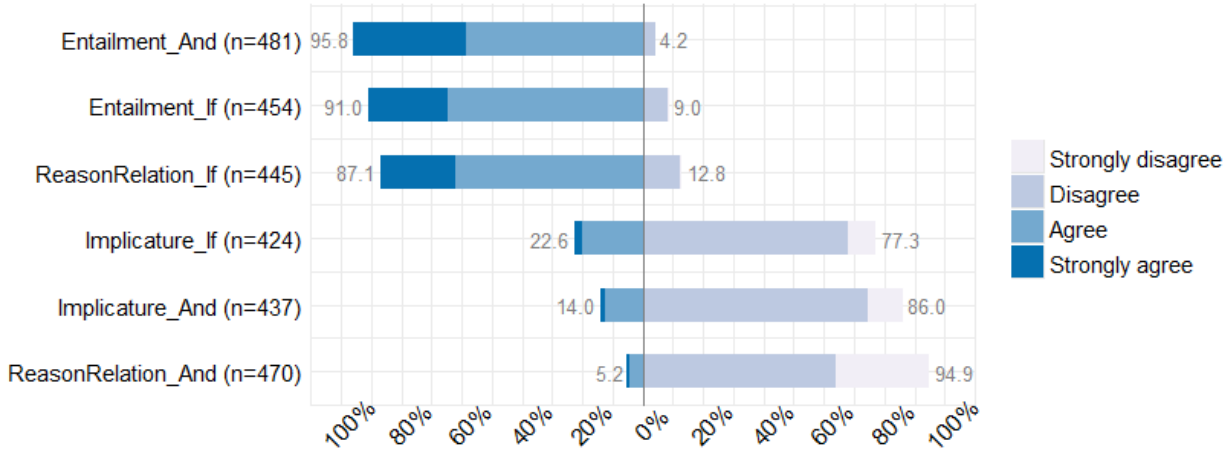
Table 4. Model Comparison

	LOOIC	ΔLOOIC	SE	WAIC	Weight
M1	3323.73	0	--	3313.2	0.9992
M2	3341.44	17.71	7.17	3329.2	0.0003
M3	3340.82	17.09	7.30	3328.6	0.0004

Note. LOOIC = leave-one-out cross-validation information criterion. WAIC = Watanabe-Akaike information criterion. Weight = Akaike weight of WAIC.

The information criteria clearly favour M1. Figure 2 plots its posterior predictions.

Figure 2. Posterior Predictions based on M1



Note. The perceived degree of contradiction of the cancellation speech act was measured on a scale from 'strongly disagree' (1), 'disagree' (2), 'neutral' (3), 'agree' (4), and 'strongly agree' for Louis' attribution of a contradiction to Samuel. The plot represents the predicted posterior probability of new participants selecting one of the displayed categories, given that they do not select 'Neutral'. For each cancellation-sentence type, the number of samples drawn from the posterior distribution is shown.

Moreover, as shown in Figure 2, there is an interesting interaction in the data whereby attempting to cancel a commitment to the reason relation is viewed as just as contradictory as attempting to cancel a commitment to an entailment, for conditionals. In contrast, for conjunctions attempts to cancel a commitment to the reason relation is viewed as less contradictory than attempts to cancel a commitment to a conversational implicature.

As a manipulation check, it can be observed across sentences that the participants clearly distinguish between attempts to cancel a commitment to entailments and conversational implicatures. While the participants agree that Samuel is contradicting himself, when attempting to cancel a commitment to an entailment, they disagree when he is attempting to cancel a commitment to a conversational implicature ($b_{\text{Implicature}}$ = -4.12, 95%-CI [-5.04, -3.23], BF_{H0H1} = $-1.44 * 10^{-38} \approx 0$).[7] The clear preference in favour of M1 (see Table 4) reflects the fact that while

[7] Note that the slightly negative BF is probably due to a minor imprecision when estimating extremely small numbers around zero in R.

there is no main effect for the Sentence factor (b_{If} = -0.44, 95%-CI [-0.90, 0.05], BF_{H0H1} = 3.93), the Sentence factor is involved in an interaction with the Cancellation Type factor. Indeed, the evidence in favour of higher ratings for attempts to cancel reason relations with conditionals than with conjunctions is very strong on conventional standards for interpreting Bayes Factors ($b_{\_ReasonRelation:If}$ = 5.02, 95%-CI [3.91, 6.17], $BF_{H0H1} = 1.28 * 10^{-16} \approx 0$).

For the positive relevance conditionals, a similar mixed effects ordinal regression model was fitted to the data, and the same data pattern was found as for irrelevance conditionals: while the participants tended to agree that Samuel was contradicting himself when attempting to cancel a commitment to an entailment, they disagreed when he attempted to cancel a commitment to a conversational implicature ($b_{Implicature}$ = -3.35, 95%-CI [-4.41, -2.34], $BF_{H0H1} = -1.89 * 10^{-18} \approx 0$). In contrast, moderate evidence in favour of the H_0 that entailment and reason relations did not differ could be obtained ($b_{ReasonRelation}$ = 0.40, 95%-CI [-0.19, 0.99], BF_{H0H1} = 6.83), with the reason relation cancellations, in fact, being rated slightly more contradicting than the entailment cancellations.

Discussion

For the control items it was found that the cancellation of an entailment was seen as more contradictory than the cancellation of a scalar implicature. While the control items concerned atomic sentences, our results indicate that this effect generalizes to the cancellation of entailments and scalar implicates occurring in conjunctions and conditionals. Consequently, we are able to use the perceived degree of contradiction of cancellation of scalar implicatures and entailments as two baselines that allow us to diagnose whether the cancellation of a reason relation suggested by conjunctions and conditionals is more like the cancellation of a conversational implicature or the cancellation of an entailment. Across conditions, it was found

that the cancellation of a reason relation of conditionals is viewed as more like cancelling an entailment than cancelling a commitment to a conversational implicature. In contrast, the reverse pattern was found for conjunctions, where it was seen as *less* contradictory to cancel a reason-relation reading than to cancel a conversational implicature. The evidence thus indicates that while both conditionals and conjunctions have reason-relation readings, the content components that contribute to them are substantially different. For conjunctions, our results indicate that the reason-relation reading has the fingerprints of a conversational implicature. This finding goes against Carston (1993), who holds in relation to the enrichment of conjunctions beyond logical conjunction that they "are instances of pragmatically derived content which contributes to the proposition expressed by the utterance, to 'what is said' in Gricean terms. That is, they are not implicatures" (p. 30).[8] And it, moreover, directly contradicts the statement on conditionals by Johnson-Laird and Byrne (2002) quoted above. In contrast, for conditionals the results suggest that the reason-relation reading does not originate in a conversational implicature, since the content component that contributes to the reason-relation reading of conditionals is not cancellable without giving rise to a contradiction.

Now it may be argued that in Johnson-Laird and Byrne's (2002) example, it seems perfectly acceptable to say "If there was a circle on the board, then there was a triangle on the board, though there was no relation, connection, or constraint, between the two—they merely happened to co-occur". Similarly, a reviewer points out that in an example like the following, we have a conditional without a connection between the antecedent and the consequent:

---

[8] However, Robyn Carston (p.c.) points out that our data for conjunctions would be compatible with other types of cancellable pragmatic inferences like those discussed in Relevance Theory (Carston, 2002). If such an explanation is viable, predictions and alternative diagnostic tests would need to be formulated for future empirical work.

Detective interviewing shopkeeper:

D: We need to know what Mr. Smith bought today, can you help us out?

S: I'm sorry, I didn't find out about any customers' names today.

D: Well, he was carrying a large polka-dotted umbrella.

S: If he carried a polka-dotted umbrella, then he bought a gold watch.

We acknowledge that there may be no relation between < there was a circle on the board; there was a triangle on the board >, nor between < he carried a polka-dotted umbrella; he bought a gold watch > at the type-level, because in general, propositions like the ones listed in these pairs are not probabilistically related. Still, at the token level, for the specific contexts in which these conditionals are used, there is a relation. If indeed 'there was a circle on the board' happens to co-occur with 'there was a triangle on the board', then there is a correlation, or a probabilistic dependency, which in fact can be used to make predictions for that specific context. What this tells us is that we need to carefully distinguish between whether the cancellation of the commitment to a connection is performed at the type or the token level. Moreover, these examples suggest that sometimes further contextual information may be needed for identifying the reason relation conveyed by indicative conditionals. For in this particular context, the proposition "Mr. Smith carried a polka-dotted umbrella" is indeed a reason for believing that Mr. Smith bought a gold watch, since it raises its probability.

Alternative Explanations

One of our reviewers suggests that implicatures vary in strength, with some being more cancellable than others (and potentially having to be cancelled in different ways), and cites the data of van Tiel et al. (2016) as evidence. Van Tiel et al. explored whether different lexical scales

give rise to scalar implicatures at different rates, and found a considerable range, from 100% of participants taking 'cheap' to implicate 'not free' to only 4% taking 'content' to implicate 'not happy'. The reviewer, we take it, uses the rate at which implicatures arise as a proxy for their strength and, inversely, cancellability.

It is an intriguing possibility that, although implicatures have classically been taken to be cancellable, some are strong enough to resist cancellation, or may require other wordings for cancellation to be effective. However, it should be kept in mind that our wording effectively cancelled the reason-relation reading of conjunctions. At first sight, it is not clear why this wording would work with conjunctions and not with conditionals that otherwise express the same content, unless there is some difference in the status of the reason-relation reading between conjunctions and conditionals. That difference, we suggest, is precisely that the reason-relation reading is a conversational implicature for conjunctions but not for conditionals.

Yet, we welcome future research comparing the reason-relation reading with a broader class of scalar implicatures that implements various wordings for cancellation. We question, though, whether van Tiel et al.'s (2016) data are sufficient to suggest that implicatures can resist cancellation. Even when scalar implicatures arise at a rate of 100%, they are intuitively cancellable. For instance, the following sentence does not seem a contradiction: 'The course is certainly cheap and may even be free'. And although 96% of van Tiel et al.'s participants took 'some' to implicate 'not all', this implicature seems readily cancellable, as in 'some - in fact, all', as, indeed, is shown by the data on the control items for Experiment 1. The rate with which scalar implicatures arise may then not be a good proxy for degree of cancellability. In the absence of countervailing evidence, we therefore retain our conclusion that whereas the reason

relation reading of conjunction could be produced by a conversational implicature, such a hypothesis is not viable for the reason-relation reading of conditionals.

A final objection is that, in Experiment 1, the contexts establish the truth of the antecedent. For instance, one item specifies that Mark is wearing socks, and continues "And if Mark is wearing socks, then it is possible that his TV will work". It is widely assumed that, for a conditional to be assertable, the truth of its antecedent must not already be settled. On a Gricean (1989) view, for instance, asserting a conditional 'if A, C' is infelicitous if the speaker knows something stronger - for instance, that A and C are both true. The infelicity supposedly arises because the speaker should respect the Maxim of Quantity and asserting something stronger - for instance 'A and C'. We might question, then, whether the present data will generalize to cases when the antecedent is not a matter of fact.

In response, we point out that existing data cast doubt on this standard assumption. In a series of studies, Krzyżanowska, Collins, and Hahn (2017b, 2018) have explored the assertability of conditionals in different contexts. It was found that conditionals can, in fact, be assertable, and not reliably less assertable than conjunctions, even when the component clauses (i.e. the antecedent and consequent) are known to be true.[9] Krzyżanowska, Collins, and Hahn's data challenge, then, the argument for why the present results might not generalize to cases when the antecedent is not known to be true. Nonetheless, empirical tests of the generalizability of our results are desirable, and desirable, too, would be further studies that directly address the range of alternative accounts suggested by the reviewer.

Experiment 2: Presuppositions

[9] A reliable difference in assertability emerged only when there was no connection between clauses (i.e. when they were irrelevant).

Having considered and rejected conversational implicatures, we turn, now, to presuppositions. Conceptually, we can think of presupposition as the marking of information that speakers take for granted when performing a speech act (Beaver & Geurts, 2014). That is, by making an assertion, the speaker acts as if the presuppositions are already an uncontroversial part of the common ground that the speaker shares with his or her interlocutors (Potts, 2015). For example, in the sentence 'Peter has stopped smoking' the word 'stopped' triggers the presupposition that Peter previously smoked. In the sentence, 'The fête was opened by the Duke of Oxford' the phrase 'the Duke of Oxford' triggers the presupposition that there is a (unique) Duke of Oxford.

The following example from Over et al. (2007) shows that the reason-relation reading of conditionals could be considered a presupposition as well:

> Consider for example:
>
> (6) If you take extra vitamin C, then your cold will be gone in three days.
>
> In most contexts, asserting (6) would be misleading, and very bad advice, if extra vitamin C was not a causal factor raising the probability that the cold will be gone in 3 days. The argument would be that there is often a pragmatic implicature when a conditional like (6) is asserted: that not taking extra vitamin C will make it probable that the cold will last longer than 3 days (p. 92)

To be sure, Over et al. (2007) here treat the example as indicating a conversational implicature, as noted above. However, one could argue that the speaker makes the assumption of a causal relation between vitamin C and getting rid of a cold as a presupposition of his or her assertion being a meaningful utterance. In addition, a presupposition failure hypothesis could be motivated by the intuition that indicative conditionals take a reason relation for granted. As Kadmon (2001:

14) says: "There is no better proof that a sentence S presupposes a proposition B than our intuition that B is 'taken for granted' and is a precondition for felicitous use of S". For these reasons, we will consider an explanation in terms of presuppositions.[10]

A sentence containing a presupposition failure has traditionally been treated as either introducing a truth-value gap or being uniformly false (von Fintel, 2004). Since the traditional semantic framework was formulated in terms of truth conditions, presupposition failures are usually conceptualized in terms of their influence on truth evaluations. However, much has happened in the field of formal semantics since Russell (1905) and Strawson (1950) had their famous debates over presupposition failures. Since many contemporary developments explicate semantic content in terms of probability distributions (Yalcin, 2012; Lassiter, 2012; Moss, 2015), it seems natural to generalize the notion of a presupposition failure to a probabilistic context as well. That is, just as a sentence may carry presuppositions for a sensible truth value assignment (such as that the entity talked about actually exists), so sentences could carry presuppositions for a coherent probability evaluation. In particular, it could be conjectured that what the Relevance Effect really shows is that indicative conditionals have the condition of Positive Relevance as a presupposition for a coherent probability assignment.

The goal of Experiment 2 is to find out whether such a linguistic phenomenon plays a role in the participants' probability assignments underlying the Relevance Effect. One of the most characteristic properties of presuppositions is their projection behavior (Karttunen, 1973).

---

[10] However, while the reason relations expressed by indicative conditionals will often be taken for granted as part of the common ground, there are argumentative uses of conditionals to introduce new reason relations in discussions, where which reason relations to accept itself becomes the content at-issue. This is, however, an issue that we will return to in the General Discussion, when considering what bearing our experimental results have for argumentation with indicative conditionals.

Projection occurs when (1) expressions are embedded under operators to form more complex expressions and (2) the presuppositions of the simpler expression are inherited by the complex expression. To see projection in action, compare the following sentences:

The Danish pope is blue-eyed.

The Danish pope is not blue-eyed.

Both sentences presuppose that there is a Danish pope. Here, we say that the presupposition projects - is constant - under negation. Presuppositions are not affected by embedding under a range of logical operators (e.g. negation, modal operators) which alter the semantic entailments of the sentences in which they occur. Accordingly, 'the family of sentences test' is one of the main diagnostics for presuppositions (Chierchia & McConnell-Ginet, 1990). In this test, it is probed whether a conjectured presupposition survives under embedding in negation, interrogation operators, as the antecedent of a conditional, and when it is placed under a possibility modal (Kadmon, 2001). For some candidate content to be a presupposition, it is a necessary condition that the candidate survives embedding under these semantic operators, but it is not a sufficient condition, since other types of content also exist which can project across them (Potts, 2015).

Note, however, that presuppositions do not always project (Gazdar, 1979; Heim, 1983). For instance, while "Peter didn't stop smoking" carries the presupposition that Peter smoked in the past, this presupposition is blocked in "Peter didn't stop smoking. He never smoked!" (Xue and Onea, 2011). But it is possible that this shows not so much that presuppositions are not characterized by their projection behavior, but rather that they are defeasible and can be cancelled, when they are embedded under other operators (Beaver and Geurts, 2011).

Consequently, to test whether the Relevance Effect is an indicator of a presupposition failure, we propose in Experiment 2 to investigate its projection behavior by embedding with irrelevant clauses under negation operators.

The Negation Task

The main purpose of the Negation Task is to test whether the relevance effect is due to a presupposition failure. One of the central characteristics of presuppositions is that they project under negation (and other embeddings). The notion of a presupposition was introduced within a truth-conditional framework, but the idea can be generalized to probabilistic content and used to account for the Relevance Effect, if the positive relevance constraint ($\Delta P > 0$) is a presupposition of a coherent probability assignment to 'if A then C'. Since ¬φ shares the same presuppositions as φ, the Relevance Effect is conjectured to be a probabilistic presupposition failure if the Relevance Effect applies equally to P(if A, then C) and P(¬(if A, then C)). That is, the same low probability assignments to P(if A, then C) in the Irrelevance condition should then be seen for the probability assignments of P(¬(if A, then C)) in the Irrelevance condition, whereas P(if A, then C) and P(¬(if A, then C)) should approximately sum to one for the Positive Relevance condition.

Alternatively, the negation operator might interact with the reason-relation reading in a way that one would expect of semantic content (although in this case it would be a case of probabilistic, semantic content). In the Irrelevance condition, both P(if A, then C) and P(if A, then ¬C) would receive low probability assignments, but P(¬(if A, then C)) may receive a high probability assignment, because [¬(if A, then C)] denies that A is a reason for C in the Irrelevance condition, whereas [if A, then C] says that A is a reason for C, and [if A, then ¬C] says that A is a reason against C. If one takes such a pattern of results together with the

dissociations between the strong effect of relevance on probability assignments and the lack of effect of relevance on truth value assignments that was found in Skovgaard-Olsen et al. (2017), one might conclude that the reason-relation reading may not affect truth conditions, but is part of the probabilistic semantic content of conditionals. For instance, one could view the reason-relation reading as a conventional implicature, which is only tapped into through probability and acceptability evaluations.

It is possible that conventional implicatures also project across embeddings under logical operators when the participants are asked for truth evaluations (Potts, 2007). But if conventional implicatures are directly tapped into by probability assignments, then they should interact with these logical operators for probability evaluations. In contrast, if probabilistic presuppositions are conditions for a coherent probability assignment to φ, then they will also be conditions for a coherent probability assignment to ¬φ, and thus project past the negation operator.

Previous Work on Negated Conditionals

Although our main interest in negations is due to their diagnostic power for determining whether the Relevance Effect is due to a presupposition failure, it is necessary to consider briefly how the Suppositional Theory and Mental Model Theory have dealt with negations of conditionals, in the interest of investigating whether these semantic theories can predict our findings.

The literature on conditionals and negation has focused on distinguishing between suppositional and mental models theories of the conditional. The Suppositional Theory makes straightforward predictions. As we have seen, the theory predicts that people judge the probability of the affirmative conditional ‘If A, then C’ to be P(C|A). It also predicts that people

judge the probability of a conditional wide-scope negation to be P(¬C|A), based on the Negation Principle (see below), and that both probabilities sum to unity. These predictions are taken to be part of the core of the Suppositional Theory. Indeed, the Negation Principle has been called a litmus test of suppositional theories as semantic theories of the conditional (Handley, Evans, & Thompson, 2006).

WIDE-SCOPE NEGATION:	¬(if A, then C)
NARROW-SCOPE NEGATION:	if A, then ¬C
NEGATION PRINCIPLE:	¬(if A, then C) <=> if A, then ¬C
	P(¬(if A, then C)) = P(if A, then ¬C)

Mental Model Theory makes more complex predictions (see, e.g., Espino and Byrne, 2012; Khemlani, Orenes, & Johnson-Laird, 2012). On the Mental Models account, an affirmative utterance is represented by a set of one or more possibilities represented by mental models; a negative utterance is represented by the complement of that set. The negation of the conditional depends on how people interpret the affirmative conditional. One possible interpretation is the so-called initial model (*A, C*), the negation of which amounts to the negation of the conjunction. But another interpretation is the fully explicit model, which contains all possibilities other than *A,* ¬*C*. In this case, the negation amounts to the conjunctive conclusion '*A* and ¬*C*' (Espino & Byrne, 2012). Yet, following the latest developments in Khemlani, Byrne, and Johnson-Laird (2016), where 'if A, then C' is thought of as expressing the conjunction ['A, C is possible' ∧ '¬A, C is possible' ∧ '¬A, ¬C is possible'], the negation of the conditional would have to express a disjunction: [¬(A, C is possible) ∨ ¬(¬A, C is possible) ∨ ¬(¬A, ¬C is possible)].

Experimental data show a complex picture. To support their account, suppositional theorists can point to evidence that people are reluctant to draw decisive conclusions from the

negation of a conditional (Handley, Evans, & Thompson, 2006). This evidence seems to suggest that people eschew the conjunctive conclusion 'A and ¬C' (Handley, Evans, & Thompson, 2006). However, participants frequently endorse the conclusion 'If A, then ¬C' (Espino & Byrne, 2012; Khemlani, Orenes, & Johnson-Laird, 2014).[11] More problematically, when participants are given more options, they also endorse 'If ¬A, then C' responses, which do not follow from the Suppositional Theory, as well as the conjunctive responses 'A and ¬C' and '¬A and C' (Espino & Byrne, 2012).

Since the Negation Principle has been used as a litmus test for the Suppositional Theory, an additional aim of Experiment 2 is to test whether it holds across relevance levels.

Method

Participants

Like Experiment 1, Experiment 2 was conducted over the Internet using Mechanical Turk and sampling from USA, UK, Canada, and Australia. 105 people participated in the experiment in exchange for a small payment. The same exclusion criteria were applied as in Experiment 1. The final sample consisted of 67 participants. Mean age was 41.3 years, ranging from 23 to 71 years; 41.8 % of the participants were male; 68.7 % indicated that the highest level of education that they had completed was an undergraduate degree or higher. The sample differed only minimally on the demographic variables above before and after applying the exclusion criteria.

Design

[11] Egré and Politzer (2013) reconciled these data sets somewhat by arguing that the negation of a conditional is, at base, understood as 'If A then possibly not C'. They then explicated contextual factors that modify this reading to recover stronger conjunctive and conditional responses.

Experiment 2 implemented a within-subject design with the following factors: Relevance (with two levels: Positive Relevance, Irrelevance), and Priors (with four levels: HH, HL, LH, LL, meaning, for example, that P(A) = low and P(C) = high for LH).

Materials and Procedure

Before beginning the study, the participants saw two control items in random order. For each, they were asked to assign a probability on a scale from 0 to 100% to a categorical sentence and to its negation, with an existential presupposition failure, using a slider. For each, the participants were randomly assigned to narrow and wide scope negations. An example of the the categorical control items and their negations is as follows:

Affirmative: The queen of the USA is in her mid-thirties.

Narrow scope: The queen of the USA is NOT in her mid-thirties.

Wide scope: It is NOT the case that the queen of the USA is in her mid-thirties.

For the main study, the 8 within-subject conditions were randomly assigned to 8 different scenarios from to the same pool of 12 scenarios from Experiment 1, for each participant anew. Similarly to Experiment 1, Experiment 2 was split into 8 blocks, one for each within-subject condition, with the same type of randomization structure as in Experiment 1.

In the context of the Mark scenario from Table 2, the participants might then be asked to make the following conditional probability judgments in the Positive Relevance HH condition:

Suppose that Mark presses the on switch on his TV.

Under this assumption, how probable is the following statement on a scale from 0 to 100%:

Mark's TV will be turned on [/Mark's TV will NOT be turned on].

In addition, the participants were asked to assign probabilities to conditional statements and their narrow and wide scope negations:

IF Mark presses the on switch on his TV, THEN his TV will be turned on.

IF Mark presses the on switch on his TV, THEN his TV will NOT be turned on.

It is NOT the case that IF Mark presses the on switch on his TV, THEN his TV will be turned on.

Results

As with Experiment 1, the within-subject design required analysis with linear mixed-effects models. As before, three models were fitted to the data using the package `brms` with weakly informative priors for mixed-effects models in Bayesian statistics with crossed random effects for intercepts and slopes by participants and by scenarios (Bürkner, 2017). Since the dependent variable consisted of continuous proportions containing zeros and ones, a zero-or-one inflated beta likelihood function was used (Ospina & Ferrari, 2012). In addition to Relevance (Positive Relevance, Irrelevance), the Type of Dependent Variable was used as a factor in the models. The types were as follows:

Affirm = P(If A, then C)

Wide = P(¬(If A, then C))

Narrow = P(If A, then ¬ C)

Of the models fitted, M4 included the 'DV Type' factor, Relevance, and their interaction. M5 was like M4, but removed the interaction. M6 was like M5 but without a main effect for the Relevance factor. Table 5 reports the performance of these models as quantified by the leave-one-out cross validation criterion and the WAIC information criterion.

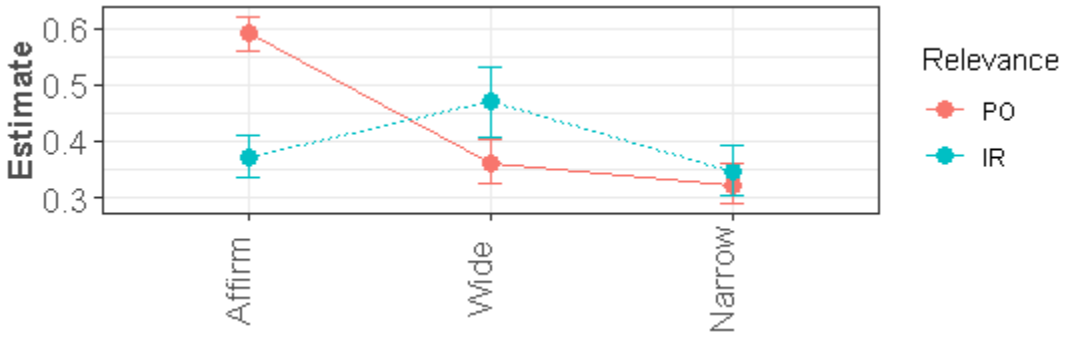
Table 5. Model Comparison

	LOOIC	ΔLOOIC	SE	WAIC	Weight
M4	1575.43	0	--	1572.8	0.922
M5	1583.80	8.36	6.14	1580.1	0.024
M6	1582.00	6.57	7.94	1578.4	0.054

Note. LOOIC = leave-one-out cross-validation information criterion. WAIC = Watanabe-Akaike information criterion. Weight = Akaike weight of WAIC.

The information criteria clearly converge on treating M4 as the winning model. This preference for M4 reflects the fact that there is a strong interaction effect in the data, which is displayed in Figure 3 through the cross-over of the lines for Affirm and Wide.

Figure 3. Posterior Mean Estimates for M4



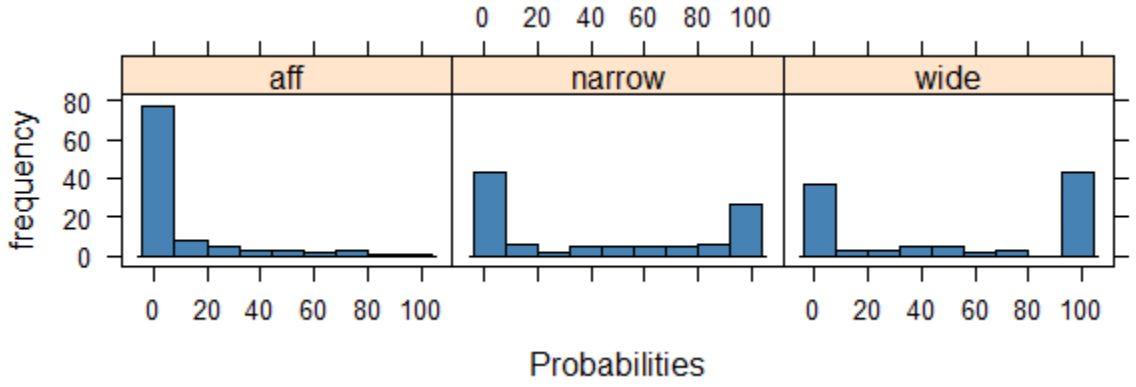
Note. 'PO' = Positive Relevance; 'IR' = Irrelevance.

When averaging across the levels of the Relevance factor, Narrow was evaluated below Affirm (b_{Narrow} = -1.43, 95%-CI [-1.68, -1.18], $BF_{H0H1} = -1.43*10^{-38} \approx 0$), and Wide was rated below Affirm (b_{Wide} = -1.21, 95%-CI [-1.46, -0.96], $BF_{H0H1} = 3.14*10^{-182} \approx 0$). And when averaging across the levels of the DV Type factor, the effect of Irrelevance was to suppress the ratings ($b_{Irrelevance}$ = -1.15, 95%-CI [-1.40, -0.91], $BF_{H0H1} = 8.17*10^{-23} \approx 0$), which is in particular visible in the large drop in the ratings of Affirm in the Irrelevance condition. However, the effect of the interaction is to substantially raise the ratings of Wide in the Irrelevance condition ($b_{Irrelevance:Wide}$ = 1.72, 95%-CI [1.29, 2.14], $BF_{H0H1} = 2.06*10^{-16} \approx 0$).

Control Items

The control items served the function of securing construct validity. When attempting to draw conclusions about whether probability assignments to conditionals across relevance levels involving negation operators are diagnostic of a presupposition failure, we need to ensure that our interpretation of a data pattern in probability ratings is actually representative for presupposition failures. To do this, we employ items that represent classical cases of presupposition failures and determine empirically how their probability ratings depend on the presence and absence of negation operators.

Figure 4. Histogram for Control Items



Note. 'wide' = wide scope negation; 'aff' = affirmativ (no negation); 'narrow' = narrow scope negation.

Since the participants differed as extremely as they could in reaction to the control items in the presence of negations, we ran a separate analysis to investigate whether the group-level performance of the participants with respect to conditionals likewise disguises such a marked individual variability.

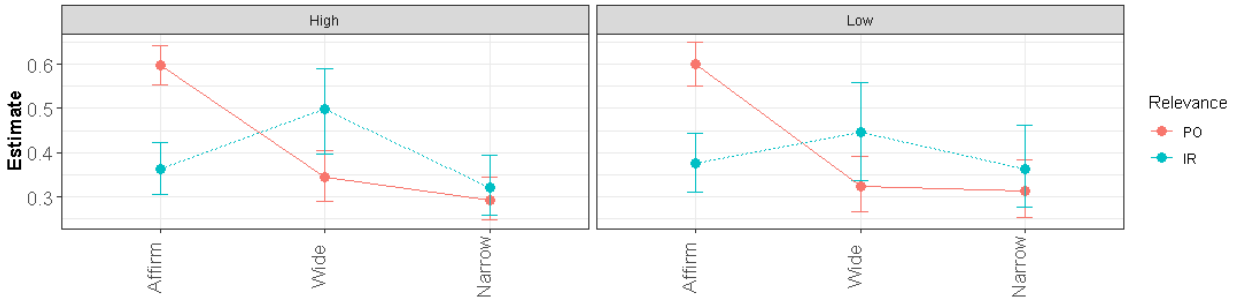
Individual Differences in Response to the Control Items

A group of 20 participants with the lowest responses (*Mdn* = 2) to the wide scope control item, and a group of 20 participants with the highest responses (*Mdn* = 82) to the wide scope control

items, were formed. It was found that the High group also assigns significantly higher probabilities to the narrow negation control item than the Low group, W = 122, p = .017, r = -.53, exact Wilcoxon rank sum test.

However, for both groups the same data pattern reported on the group level in Figure 3 was found (see Figure 5 below). This in turn indicates that whatever individual differences in the probability assignments to presupposition failures are found in the control items, they are not matched by the participants' probability assignments to conditionals in the Irrelevance condition. For both groups, there is little overlap between their behavior with respect to the control items and the conditionals in the Irrelevance condition, as shown in Figure 5 based on a mixed linear model like the previous, which, however, included the Group factor along with its interactions:

Figure 5. Group Low and High in Comparison



Note. Plots show the posterior means for the Low group and the High group. The groups were formed based on the participants' responses to the control items with wide scope negation. 'PO' = Positive Relevance; 'IR' = Irrelevance.

Probabilistic Coherence

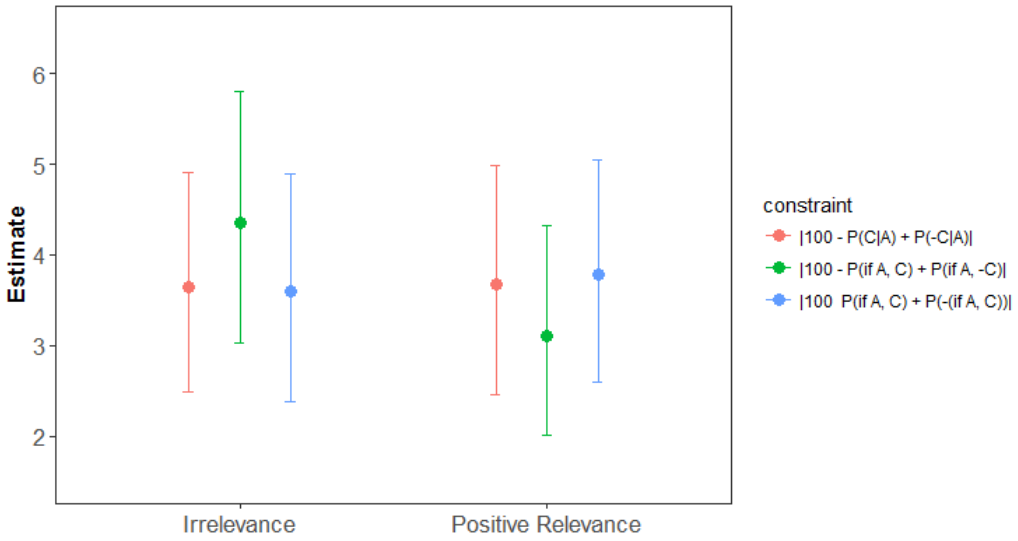
As part of Experiment 2, we investigated the participants' probabilistic coherence, that is, whether complementary pairs of participants' probability assignments sum to 100%. We did so on the grounds that, if reason relations are supposed to be made part of the semantic, probabilistic content of indicative conditionals, then probabilistic coherence makes up a natural requirement. In Appendix 1, we present analyses based on linear mixed-effects models in classical statistics which indicate violations of probabilistic coherence in the Irrelevance

condition due to the impact of an influential group of outliers, which violates additively strongly in both directions (either by having values of complementary pairs summing to 0 or 200 in the Irrelevance condition). For this reason, we here follow Kruschke's (2014) recommendation of conducting a robust regression analysis with a t-distribution as the likelihood function, which is less sensitive to the influence of outliners. As before, we did the analyses using mixed-effects models in Bayesian statistics with crossed random effects for intercepts and slopes by participants and by scenarios (Bürkner, 2017). The linear mixed-effects models treated the participants' probability judgments as a function of Relevance (Positive Relevance, Irrelevance) and Constraint Type (see below), and their interaction, which was allowed to vary across participants and scenarios. Three such models were contrasted: (M7) modeled the rating as a function of the factor Relevance, the factor Constraint Type, and their interaction, (M8) like M7 but without the interaction, and (M9) like M8 but without a main effect for the Constraint Type factor. The Constraint Type factor had the following levels:

P(C|A) Consistency:	$y = \|100 - (P(C\|A) + P(\neg C\|A))\|$
Narrow Consistency:	$y = \|100 - (P(\text{If A, then C}) + P(\text{If A, then } \neg\text{C}))\|$
Wide Consistency:	$y = \|100 - (P(\text{If A, then C}) + P(\neg(\text{If A, then C})))\|$

That P(C|A) + P(¬C|A) = 100% across relevance conditions is a general requirement of probabilistic coherence when reasoning with conditional probabilities. That P(If A, then C) + P(¬(If A, then C)) = 100% across relevance conditions is a requirement of conditional consistency that theories of conditionals across the board should accept. In contrast, that P(If A, then C) + P(If A, then ¬C) = 100% is a requirement that theories of conditionals adopting the Negation Principle should accept.

Figure 6. Probabilistic Coherence



Note: The plot displays departures from probabilistic coherence on the three investigated measures using robust regression analysis with a t-distribution as the likelihood function.

As Figure 6 indicates, the responses of the participants were shown to be remarkably consistent once the influence of outliers was controlled through robust regression analyses. This contrasts with the analysis reported in Appendix 1, where particularly large departures from probabilistic coherence were found for the Narrow Consistency constraint in the Irrelevance condition (see Figure 6, Appendix 1).

Table 6 reports the performance of these models as quantified by the leave-one-out cross validation criterion and the WAIC information criterion.

Table 6. Model Comparison

	LOOIC	ΔLOOIC	SE	WAIC	Weight
M7	17069.63	0.66	4.36	17068.8	0.4080
M8	17072.61	3.64	0.52	17072.1	0.0819
M9	17068.97	0	--	17068.4	0.5101

Note. LOOIC = leave-one-out cross-validation information criterion. WAIC = Watanabe-Akaike information criterion. Weight = Akaike weight of WAIC.

The model comparison in Table 6 indicates a tight run between M7 and M9 with a slight preference for M9. The preference for M9 is explained by the lack of a main effect for Constraint

Type ($b_{\text{constraintP(C|A)}}$ = 0.05, 95%-CI [-0.81, 0.89], BF_{H0H1} = 7.00; $b_{\text{constraintNarrow}}$ = 0.76, 95%-CI [-0.20, 1.76], BF_{H0H1} = 1.84). Similarly, there was no main effect of Relevance ($b_{\text{PositiveRelevance}}$ = 0.18, 95%-CI [-0.62, 1.00], BF_{H0H1} = 6.63). The relative competitiveness of M7 indicates that the Constraint Type factor was involved in an interaction with the Relevance factor, however, which is also the reason why removing the interaction term in M8 results in the worst fitting model.

A closer look at the interaction indicates that while there is moderate evidence for H_0 when setting the coefficient for the P(C|A) Consistency constraint equal to zero ($b_{\text{PositiveRelevance:constraintP(C|A)}}$ = -0.15, 95%-CI [-1.23, 0.94], BF_{H0H1} = 5.23), there is anecdotal evidence against H_0 when setting the coefficient for the Narrow Consistency constraint equal to zero ($b_{\text{PositiveRelevance:constraintNarrow}}$ = -1.42, 95%-CI [-2.82, -0.16], BF_{H0H1} = 0.43).

Discussion

The results indicate that the Negation principle (¬(if A, then C) <=> if A, then ¬C) could only be maintained for the Positive Relevance condition; for the Irrelevance condition it was systematically violated.

For our diagnostic purposes, it is particularly interesting that the wide scope negation (¬(if A, then C)) was rated the highest for the Irrelevance condition. This finding indicates that, whatever content component is responsible for the low probability assignments in the Irrelevance condition (dubbed 'the Relevance Effect'), it does not project under wide scope negation. Yet surviving embeddings under the semantic operators listed in the family of sentences test is treated as a necessary condition for the targeted content to count as a presupposition (Potts, 2015). This suggests that the Relevance Effect is probably not produced by a presupposition failure (with conditionals requiring that a Positive Relevance constraint is satisfied in order to

obtain a high probability). To substantiate this interpretation, a comparison was made with the control items (discussed below). But note that we emphasize that passing the family of sentences test is a necessary, but not sufficient, condition for counting as a presupposition, because other types of content are known to have instances where they project as well. In particular, the conventional implicature 'but' expresses a contrastive relation that projects in examples like 'If John is rich, but happy, then…'. However, unlike classical presupposition triggers like 'know' and 'again', conventional implicatures like 'but' and 'therefore' are only able to pass the family of sentences test for some of its instances, it seems ('?' suggesting possible failures):

? It is not the case that John is rich but happy

? It is possible that John is rich but happy

? Is John rich but happy?

If John is rich but happy, then ….

? It is not the case that John is rich therefore he is happy

? It is possible that John is rich therefore he is happy

? Is John rich therefore he is happy?

? If John is rich therefore he is happy, then …

Arguably, in order for the embedded therefore-sentences even to be grammatical, it seems that they would need to be rewritten as 'John is rich *and therefore* he is happy', but then 'and' takes the role of being the main connective (see also McCawley, 1993, pp. 318-319).

Control Items

In the present case, our control condition introduces the rather curious finding that the participants differ as extremely as they can on a probability scale on their reaction to the control item in the presence of negations (in particular with respect to wide scope negations).

For this reason, we did a separate analysis to investigate whether the group-level performance of the participants with respect to conditionals likewise disguises such a marked individual variability. As we saw, this turns out not to be the case (since the same qualitative differences with respect to the probability assignments of the conditional at the group level is found within each group who differed in their performance on the control item, Figure 5). This in turn can be interpreted as supererogatory evidence for the above conclusion that the participants' performance with conditionals in the Irrelevance condition is not an expression of how they generally treat presupposition failures. Although both groups differed markedly in how they assign probabilities to a paradigmatic case of presupposition failures, the two groups showed the same qualitative pattern when assigning probabilities to conditionals. Furthermore, their common data pattern in relation to conditionals was not matched with respect to the control items for either group. Whatever is driving the participants' responses for the conditionals in Irrelevance conditions is in other words unlikely to be the result of how the participants process presupposition failures in general.

Moreover, looking back to the results from Experiment 1, it is worth keeping in mind that conversational implicatures and presuppositions can be denied without contradiction, whereas conventional implicatures and entailments cannot be denied without contradiction (Potts, 2007). Yet in Experiment 1 it was found that denying the reason-relation reading of conditionals is viewed as just as contradictory as denying an entailment. It has been observed, though, that the

examples with cancellable presuppositions seem to concern specifically embedded presuppositions (Beaver & Geurts, 2011), which was not what was investigated in Experiment 1.

Local Accommodation?

A reviewer suggested that the lack of projection behavior found for embedding conditionals in the Irrelevance condition under the wide-scope negation operator could be explained by a process of local accommodation, which essentially turns presuppositional content into regular entailed content. To illustrate, von Fintel (2008) discusses the following example by Heim (1983):

There is no king of France. Therefore, the king of France is not hiding in this room.

The point is that accepting the first sentence updates the common ground so that it is taken for granted that there is no king of France. Against this utterance context, it is not possible for the second sentence to presuppose that there is a king of France. Instead, the presupposition of the second sentence is cancelled by a process of local accommodation. More specifically, the context is treated as updated vacuously by an empty set consisting of the contradictory context that would be generated had 'there is a king of France' been added to the context set.

The suggestion is then that the reason why we find the highest probability ratings for the wide negated conditional in the Irrelevance condition is because such a process of local accommodation cancels the presuppositional content and blocks its projection behavior.

In response, we would like to point to a number of disanalogies between our experimental task and examples like the one above. First of all, there is no conversational context, no utterance of the conditionals as assertions, and no prior updates of the common ground in our experimental task. The participants merely see the conditional sentences and are asked to assess their

probabilities. To be completely parallel with the example above, the participants would have to be presented with assertions of the following kind by a speaker:

[A] is irrelevant for [C]. Therefore, it is not the case that if [A], then [C].

Had the participants been presented with such materials, it could be argued that local accommodation would cancel the $\Delta P > 0$ presupposition of the conditional, since the first sentence already rules out this presupposition. But given that our experimental task was different, we would still need to see it be established that local accommodation can arise for sentences such as ”If Mark is wearing socks, then his TV will work”, for which it is obvious that the two things are unrelated and where one cannot rely on the participants to modify the context such that these sentences become accepted.

Negation Principle

The results indicate that the simple relationship by which P(C|A) predicts P(if A, then C) and P(¬C|A) predicts both P(¬(if A, then C)) and P(if A, then ¬C) across relevance levels does not hold. In particular, this relationship is generally less strong for Irrelevance than for Positive Relevance, and less strong for wide scope negation (where it is almost absent) than for narrow scope negation. This in turn violates the negation principle (P(¬(if A, then C)) = P(if A, then ¬C), which was found to hold only for the Positive Relevance condition.

Moreover, the results indicate that there is an interaction with the negation operator and the relevance factor, which makes ‘¬(if A, then C)’ the most probable statement in the Irrelevance condition. That is, the low probability assignment to ‘if A, then C’ in Irrelevance is not matched by a low perceived probability of the wide scope negated conditional in Irrelevance. As already discussed, this is evidence against an interpretation in terms of a presupposition

failure which projects under wide scope negation; evidence to which the pattern of results for the control items adds further weight.

How do the present data on the Negation Principle bear on the existing literature? Our data bear most straightforwardly on the Suppositional Theory of conditionals. Recall that suppositional theorists have themselves called the Negation Principle a litmus test of the Suppositional Theory (Handley et al., 2006). Our data suggest that the Negation Principle does not hold in general, and that an important qualification is needed. Moreover, Mental Model Theory is not straightforwardly able to predict systematic differences according to relevance condition. Recall that, for Byrne and Johnson-Laird (2002), reason relations best correspond to a conversational implicature, but Experiment 1 has already cast doubt on this possibility at least for affirmative conditionals.

Probabilistic Coherence

If reason relations are supposed to be made part of the semantic, probabilistic content of indicative conditionals, then probabilistic coherence makes up a natural requirement. In our analysis, we tested three requirements of probabilistic coherence: probabilistic coherence of conditional probabilities ($P(C|A) + P(\neg C|A) = 100\%$ across relevance conditions), conditional consistency ($P(\text{If A, then C}) + P(\neg(\text{If A, then C})) = 100\%$ across relevance conditions), and a conditional consistency requirement based on the Negation Principle ($P(\text{If A, then C}) + P(\text{If A, then } \neg C) = 100\%$ across relevance conditions).

Divergences from all three requirements of probabilistic coherence were found for the Irrelevance condition, when outliers were included (see Appendix 1). The strongest divergences were found for the conditional consistency requirement based on the Negation Principle. In contrast, it was possible to find high levels of probabilistic coherence for all three constraints

across relevance levels, when the influence of outliers was controlled through robust regression techniques within a Bayesian framework (see Figure 6).

A priori we had only expected to find violations of probabilistic coherence for the Irrelevance condition, if $\Delta P > 0$ were made a presupposition of coherent probability assignment (just as the existence of what one is talking about can be made a presupposition of a truth value assignment). However, given the weight of the evidence cited above against the presupposition failure account of the Relevance Effect, the violations of probabilistic coherence in the Irrelevance condition when outliers are included need to be accounted for differently. We offer the following *post hoc* explanation. With 24 combinations of negation operators (no negation, narrow scope, wide scope), prior (HH, HL, LH, LL), and relevance levels (Positive Relevance, Irrelevance) for conditionals, and 16 combinations of conditional probabilities, it can be challenging to both pay attention to fine details in the stimulus materials and maintain consistency internally in the responses given. Accordingly, limitations in the cognitive resources invested in maintaining internal consistency in the online study may account for the violations of probabilistic coherence found in the group of outliers.

Experiment 3

Given that Experiments 1 and 2 cast doubt on the hypotheses that the Relevance Effect arises from conversational implicature or presupposition failures respectively, Experiment 3 investigated whether it arises from conventional implicature. For this purpose, Experiment 3 investigates a much-discussed property of conventional implicatures: that conventional implicatures are content that is not-at-issue.

Conventional Implicatures

Grice (1989) noticed meanings which do not seem to contribute to the truth or falsity of a sentence but which nevertheless seem conventional. He termed such meanings 'conventional implicatures'. Classic examples include words such as 'but', 'therefore' and 'even' (Potts, 2007; Valleé, 2008; Salmon, 2011; Blome-Tillmann, 2013). For instance,

'He is English but brave.'

'Even the English can be brave'

'He is English and, therefore, brave.'

In the first sentence, 'but' signals a contrast between Englishness and bravery: bravery is somehow unexpected from the English. In the second sentence, 'even' likewise signals unexpectedness, perhaps also implicating that bravery is not so exceptional after all if the English are capable of displaying it. In the third sentence, 'therefore' signals a consequence relation between Englishness and bravery.

Each example above makes a claim that is contentious. Many readers might want to deny that there is a relationship - positive or negative - between Englishness and bravery. But according to Grice (1989), a reader so minded would struggle to say that the relationship expressed makes the sentences false. This is not, of course, to say that the sentences cannot be true or false. Readers can easily deem the first and third examples false if they believe that the 'he' in question is a coward or not English (or both), and can deem the second example false if they believe all English people to be incapable of bravery. But according to Grice (1989), the relationship between Englishness and bravery does not contribute to the truth evaluation of the sentence. Accordingly, these meanings - these conveyed relationships - do not seem to be straightforwardly semantic on the typical understanding of the term. There is, nevertheless,

something conventional about these meanings, since they attach to specific words. Equally, these meanings are not straightforwardly pragmatic: they are not calculable based on the Gricean maxims, and they cannot be cancelled without contradiction.

Conventional implicatures contrast markedly with both presuppositions and conversational implicatures. Presuppositions do affect the truth evaluation of the sentences in which they occur, either by creating a truth value gap - the sentence is neither true nor false - or making the sentence uniformly false. Conversational implicatures are calculable based on the Gricean maxims, and can be cancelled without contradiction.

One clear indicator that the Relevance Effect might be produced by a conventional implicature is found in a truth-table task with relevance manipulations in Skovgaard-Olsen et al. (2017). Across two experiments, a strong dissociation was found indicating that while the reason-relation reading of conditionals clearly influences probability and acceptability evaluations, it has almost no influence on truth evaluations. These results might be interpreted as support for a conventional implicature hypothesis, according to which the reason-relation reading of indicative conditional is a conventional aspect of their meaning, which cannot be cancelled without contradiction, is not calculable by the Gricean maxims, and is not targeted by truth evaluations. According to this interpretation, the reason-relation reading of conditionals would be similar to the reason-relation reading of 'A but C' and 'A therefore C ' (which in turn suggest that A is a reason against or for C, respectively). Support for this interpretation can also be derived from the fact that Skovgaard-Olsen et al. (2017) found the same strong dissociations with respect to but- and therefore-sentences as a function of the reason-relation reading that they found for the indicative conditionals.

One property that seems to be attributed to both presuppositions and conventional implicatures in recent discussions is that they are content that is not-at-issue. It has long been observed in relation to presuppositions that they are backgrounded content which is taken for granted (Kadmon, 2001). Recently, a similar idea has come to play a major role in treatments of conventional implicatures. The backdrop is that Bach (1999) made an influential case that appositives (e.g. 'Mozart, *the famous composer*, used to live here') should be treated as a central instance of conventional implicatures, instead of the kind of examples Grice (1989) considered. This recommendation was later taken up and fully developed in Potts (2005), where a formal system was developed that treats conventional implicatures as logically and compositionally independent of the at-issue content.

However, whereas the not-at-issue content of presuppositions is thought of as backgrounded and known by the participants, Potts (2015) argues that conventional implicatures usually introduce new content that is not-at-issue. One way in which this difference shows up is that whereas it is usually not redundant to first explicitly state a presupposition and then make an utterance carrying this presupposition, it is considered redundant to explicitly assert conventional implicatures, because their content is lexically encoded (Potts, 2007). To illustrate, whereas it is redundant to say 'Mozart is a famous composer' and 'Mozart, the famous composer, used to live here', it is not considered redundant for a speaker to explicitly state that in his or her view 'there is a Danish pope' and 'the Danish pope is blue-eyed'.

A further difference is that whereas the not-at-issue content of presuppositions is thought of as cancellable in compound sentences, one cannot cancel a lexically encoded conventional implicature without contradiction (Potts, 2015).

Tonhauser (2012) collects a series of diagnostic tests for deciding whether a content component makes up content-at-issue. Central among these is that content-at-issue is targeted by direct denials ("No, ...") and acceptances ("Yes, ..."), whereas content-not-at-issue can only be denied by more indirect measures that interrupt the natural flow of the conversation like "Actually, ...", "Well, ...", "Hey, wait a minute!".

In Syrett and Koev (2014) a series of experiments was conducted with appositives based on this idea. While the authors found clear preferences for direct denials as targeting the main claim in utterances containing appositives, the evidence was more mixed when it comes to whether a preference for "Hey, wait a minute!" can be used as a sufficient condition for identifying content not-at-issue. This picture fits with the corpus analysis of natural occurrences of "Hey, wait a minute!" conducted in Potts (2008), which also did not find compelling evidence that this construction is only used with not-at-issue content like appositives. Similarly, Salmon (2011) also finds that "Hey, wait a minute!" rejections can be used with other constructions than their intended purpose. For this reason, we decided against employing this test for our tests of whether the reason relation constitutes content-at-issue of indicative conditionals. Instead, we decided to investigate which content component the participants naturally interpret direct acceptances and denials as targeting in Experiment 3, following the diagnostic tests of at-issue content in Tonhauser (2012).

Method

Participants

Like Experiment 1, Experiment 3 was conducted over the Internet using Mechanical Turk and sampling from the USA, UK, Canada, and Australia. 339 people participated in the

experiment in exchange for a small payment. The same exclusion criteria were applied as in Experiment 1. The final sample consisted of 228 participants. Mean age was 40.8 years, ranging from 19 to 98[12] years; 41.7 % of the participants were male; 73.7 % indicated that the highest level of education that they had completed was an undergraduate degree or higher. The sample differed only minimally on the demographic variables above before and after applying the exclusion criteria.

Design

Experiment 3 implemented a mixed design with the following within-subject factors: relevance (with two levels: Positive Relevance, Irrelevance), priors (with four levels: HH, HL, LH, LL, meaning, for example, that P(A) = low and P(C) = high for LH), and truth table cell (TT vs. TF). There was one between-subject factor: sentences (If then vs. Therefore).

Materials and Procedure

Because we did not want to use a scenario twice for a given participant, each participant underwent only 12 of the 16 within-subject conditions. These comprise the 8 conditions elicited by crossing relevance and prior factors, and a random 4 additional conditions from the relevance x prior subdesign. Truth-table cells (TT vs. TF) were randomly assigned to the 12 conditions, for each participant anew, so that 6 conditions were allocated with TT and 6 with TF.

Experiment 3 was split into 12 blocks with the same type of structure as in Experiment 1 (unless otherwise noted). At the beginning of the experiment, the participants were presented with items containing appositives for exploratory purposes, which are not reported here.

[12] It is doubtful whether the response '98' really should be taken at face value. The next highest was '72'.

For the experimental task the participants were instructed that they should help Pierre, a foreign exchange student, to learn English by correcting his mistakes. Their task was to indicate whenever Pierre had said something that was true and to correct him whenever he had said something that was false.

Below the scenario, the participants would see two facts that they and Pierre had learned about the scenario and a statement that Pierre had made. For instance, one participant might have seen the Mark scenario text from Table 2, continued as follows.

You and Pierre both learn that:

Mark presses the on switch on his TV.

Mark's TV will be turned on. [/Mark's TV will NOT be turned on]

Pierre: IF Mark presses the on switch on his TV, THEN his TV will be turned on.

[/Mark presses the on switch on his TV THEREFORE his TV will be turned on]

The participants were then asked whether Pierre had made a true statement and they were instructed to give their response as a forced choice between two options which varied with the condition. In this example, the participants were then given a forced-choice between:

Yes, Mark presses the on switch on his TV and his TV will be turned on vs.

Yes, Mark presses the on switch on his TV is a reason that his TV will be turned on

The remaining conditions followed a similar structure:

1. Positive Relevance TT "Yes, TT" vs. "Yes, A is a reason for C"

Forced Choice between two Reasons to accept

2. Irrelevance TF "No, TF" vs. "No, A is not a reason for C"

Forced Choice between two Reasons to reject

3. Positive Relevance TF "No, TF" vs. "Yes, A is a reason for C"

Forced Choice between Yes/No Answers + Justifications in a Conflict Case

4. Irrelevance TT “ “Yes, TT” vs. “No, A is not a reason for C”

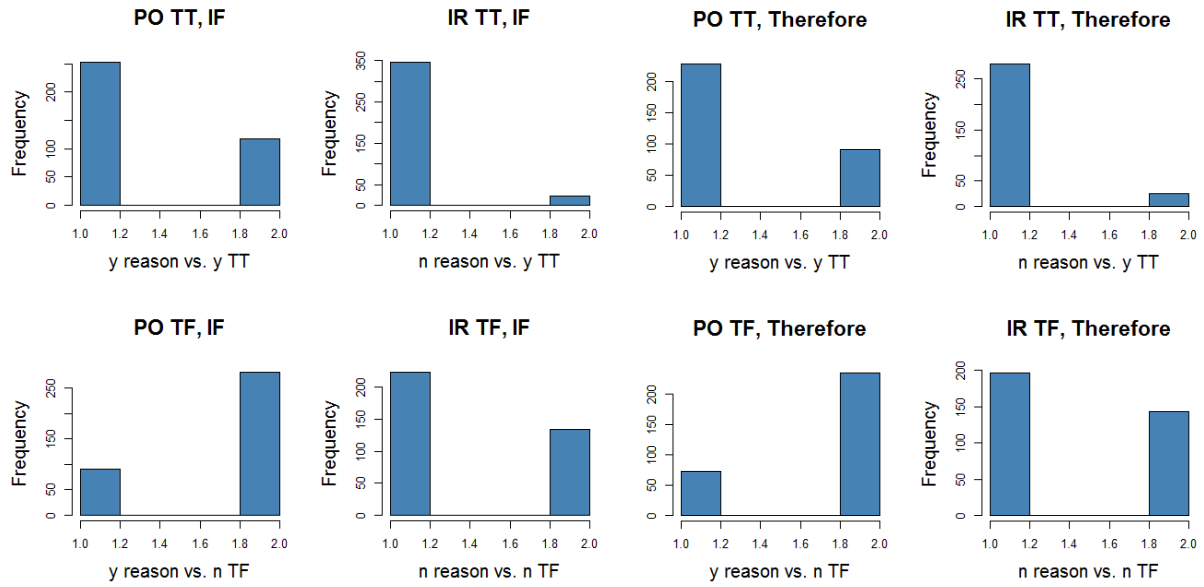
Forced Choice between Yes/No Answers + Justifications in a Conflict Case

If the reason-relation reading is content-at-issue for the participants, then the reason-relation justification should appear most attractive.

Results

As Figure 7 indicates, there was a widespread tendency to treat the reason-relation reading as content-at-issue for both indicative conditionals and therefore-sentences.

Figure 7. Histograms of Conditionals and Therefore



Note. 'y reason' = 'yes, A is a Reason for C'; 'n TF' = 'No, TF'; ‘y TT’ = ‘yes, TT’.
‘PO’ = Positive Relevance; ‘IR’ = Irrelevance.

The following constrained multi-nominal processing models were fitted to the data (Batchelder and Riefer, 1999; Skovgaard-Olsen et al., 2017):

$M_{\text{saturated}}$: model imposing no constraints. This model fits the data perfectly using one free parameter per degree of freedom provided by the data

$M_{sentence}$: model assuming that response probabilities are the same across sentences, while allowing for differences across the different levels of the Relevance factor

$M_{relevance}$: model assuming no differences across the different levels of the Relevance factor while allowing for differences across the levels of the Sentence factor

M_{full}: model assuming no differences across both the relevance and sentence levels

$M_{ad\ hoc}$: A post hoc modification of $M_{sentence}$ to allow the sentences to differ in the Irrelevance TF cell.

A hypothesis (instantiated by a constrained model) is said to be rejected when it performs worse than the unconstrained model and/or any of the competing alternative hypotheses (even after taking differences in flexibility into account via the Fisher Information Approximation, FIA). In addition, the ratio G^2/df is included since for large samples any minor deviation from model predictions can lead to a statistically significant misfit. A ratio G^2/df between 0 and 2 is considered to indicate a good fit (Skovgaard-Olsen et al., 2017).

Table 7. Model-Comparison Results: Therefore/If then

Model	G^2	df	*p*	ΔFIA	G^2/df
$M_{saturated}$	0	0	1	9.20	--
$M_{sentence}$	3.94	4	.41	0	0.94
$M_{relevance}$	318.83	4	.00	157.44	79.71
M_{full}	319.17	6	.00	151.33	53.20
$M_{ad\ hoc}$	2.37	3	.50	2.02	.79

Note. df = degrees of freedom; G^2 = goodness of fit; p = p-value; ΔFIA = difference between the model's FIA and the FIA from the best-performing model.

Based on these criteria, a model ($M_{sentence}$) that collapses the difference between conditionals and therefore sentences is the winning model, as Table 7 shows.

The main findings are: 1) for the TT cell there is a preference for treating the reason-relation reading as at-issue content across the levels of the Relevance factor and the Sentence

factor 2) in the Positive Relevance TF cell, there is a preference for treating the truth-table cell as at-issue across the levels of the Sentence factor, and 3) in the Irrelevance TF cell, there is a tendency for treating the reason-relation reading as at-issue across sentence levels.

Discussion

Taken together, the reason-relation reading turned out to be content-at-issue in the context of the present task, not only for the conditionals, but also for the therefore-sentences, which were used as a baseline representing a paradigmatic instance of Gricean conventional implicatures. Yet, interestingly, the participants' preferences for the TF conditions indicate that while a TF cell is sufficient to elicit a ‘no’ response, and thereby trump the reason relation in the Positive Relevance TF condition, the reason relation is weighted slightly higher when there are two justifications for answering ‘no’ in the Irrelevance TF condition.

In Skovgaard-Olsen, et al. (2017), it was found that there was little influence of relevance on truth-value assignments to both conditionals and 'therefore' sentences in a truth table task where the participants were instructed to calibrate the output of a computer program in the development phase. One of the main differences between the computer calibration task used there and the present Pierre task is the following: (1) in the Pierre task the sentences were produced based on known truth-table cells, whereas the sentences were produced before the truth-table cells became known in the computer calibration task, and (2) in the Pierre task a forced choice between Yes/No responses in the presence of justifications is required, whereas the computer calibration task involved a ternary assignment {True, False, Neither Nor} without justifications.

It has become popular to think about conventional implicatures as content-not-at-issue based on the influential work of Bach (1999) and Potts (2005). In the present experiments, we could not find evidence that the reason-relation reading of conditionals is content not-at-issue. However, as argued in Salmon (2011), the Bach-Potts notion of conventional implicatures, which centers around the example of non-restrictive relative clauses and appositives (E.g. "Mozart, who is a famous composer, started to play the piano at an early age"), is subtly different from the notion of conventional implicatures that figures in the work of Grice (1989), where the main examples came from sentences containing words like 'but', 'therefore', and 'even'.

These two notions of conventional implicatures are supposed to share the properties of a) not affecting the truth values of the sentences in which they occur, and b) being conventional aspects of the meaning of the sentences in which they occur, which cannot be canceled without contradiction (as opposed to the pragmatic content of conversational implicatures).

It was an assumption of our experiments that they would also share the property of being content not-at-issue. Based on the results of our present study, it appears that neither Gricean conventional implicatures (here represented prototypically by the therefore-sentences), nor conditionals have this property. Conversely, the success of the model $M_{sentence}$ implies that conditionals were responded to just like therefore-sentences in our test of at-issue content.

Experiment 4

It is possible that, in Experiment 3, participants took the truth cells provided to be evidence for or against a reason relation. In Positive Relevance items, the TF cell might be taken as evidence against Positive Relevance. For instance, although we would normally assume that Mark pressing the on switch on his TV is a reason for believing that his TV will be turned on, it might

be thought that this reason relation is undermined by learning that while Mark has pressed the on switch on his TV in fact his TV is not turned on. A similar problem arises in Irrelevance items, where the TT and TF conditions might be taken as evidence of reason relations. If participants did, indeed, reason this way, this reasoning would undermine the relevance relation manipulation through the scenarios based on the participants' background knowledge. Experiment 4 tests this alternative hypothesis, and is a control study for Experiment 3. Its purpose is to ensure that the TF condition does not undermine the Positive Relevance manipulation, and that the TT and TF conditions do not undermine the Irrelevance manipulation.

Method

Participants

Like Experiment 3, Experiment 4 was conducted over the Internet using Mechanical Turk and sampling from USA, UK, Canada, and Australia. 250 people participated in the experiment in exchange for a small payment. The same exclusion criteria were applied as in Experiment 3. The final sample consisted of 155 participants. Mean age was 36.9 years, ranging from 21 to 73 years; 52.9 % of the participants were male; 64.5 % indicated that the highest level of education that they had completed was an undergraduate degree or higher. The sample differed only minimally on the demographic variables above before and after applying the exclusion criteria.

Design

Experiment 4 had the same experimental design as Experiment 3.

Materials and Procedure

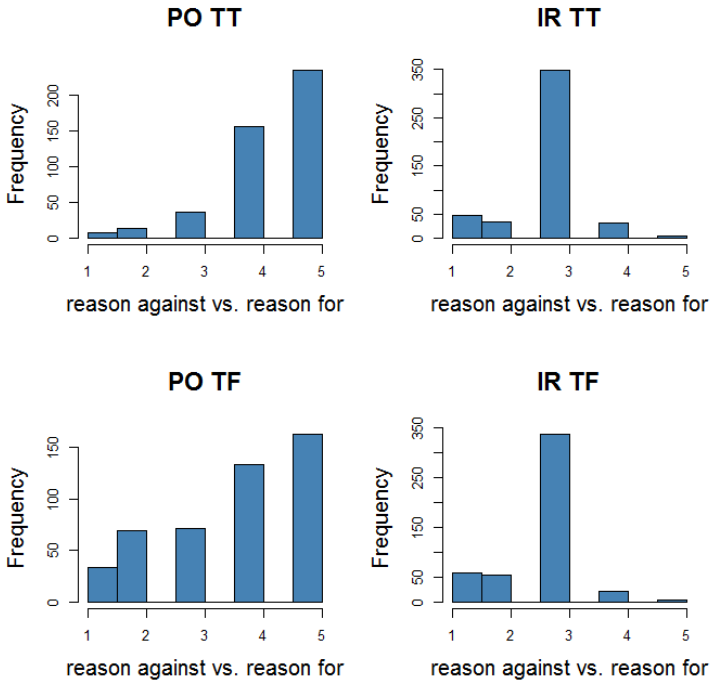
Experiment 4 followed the same procedure as Experiment 3. The only difference was that instead of making judgments of direct rejections and affirmations, the participants were now asked to rate the extent to which the antecedent of Pierre's statement was a reason for/against the

consequent on a five-point scale {strong reason against; reason against; neither for nor against; reason for; strong reason for}. These two randomly ordered pages differed on whether the participants were presented with the TT or TF cell as the facts that they and Pierre learned about.

Results

Participants assessed on a five-point scale the extent to which the first sentence in our Positive Relevance and Irrelevance conditions was a reason for/against the second. In Figure 8 the histograms are displayed. In Positive Relevance TT, the participants' reason-relation rating was 'strong reason for' (*Mdn* = 5) on the group level; in Positive Relevance TF it was 'reason for' (*Mdn* = 4). The reason relation in the Irrelevance condition was not affected by the presence of truth table cells and was in both cases assessed as 'neither for nor against' (Irrelevance TT *Mdn* = 3; Irrelevance TF *Mdn* = 3).

Figure 8. Histogram of Reason-Relation Assessments



Note. 1 = strong reason against; 2 = reason against; 3 = neither for nor against; 4 = reason for; 5 = strong reason for. 'PO' = Positive Relevance; 'IR' = Irrelevance.

Linear Mixed Models

To analyze the effect of the truth table cell on the reason-relation assessments, linear mixed-models were fitted to the participants' responses. Like in Experiment 1 where a Likert scale was also used, the responses were modelled as generated by thresholds set on a latent continuous scale with a cumulative likelihood function and a logit link function (Bürkner & Vuorre, 2018). The models included the following fixed effects: (M10) modelled the rating as a function of the Item factor (TT vs. TF), the Relevance factor (Positive Relevance vs. Irrelevance), and their interaction, (M11) like M10 but without the interaction, and (M12) like M11 but without the main effect for the Relevance factor.

Table 8 reports the performance of these models as quantified by the leave-one-out cross validation criterion and the WAIC information criterion.

Table 8. Model Comparison

	LOOIC	ΔLOOIC	SE	WAIC	Weight
M10	4024.12	0	--	4002.1	0.9939
M11	4042.36	18.24	6.05	4019.0	0.0002
M12	4035.24	11.12	4.66	4012.4	0.0059

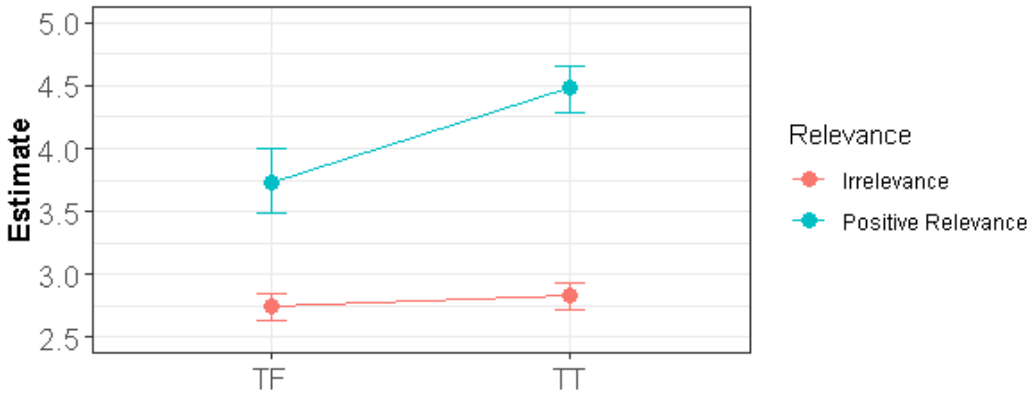
Note. LOOIC = leave-one-out cross-validation information criterion. WAIC = Watanabe-Akaike information criterion. Weight = Akaike weight of WAIC.

The information criteria clearly favour M10. This preference for M10 reflects the fact that there was a strong interaction between the Item and Relevance factors ($b_{\text{ItemTT:PositiveRelevance}}$ = 1.69, 95%-CI [1.08, 2.31], $BF_{H0H1} = 4.60*10^{-11} \approx 0$). In addition, a strong main effect for the Relevance factor was found ($b_{\text{PositiveRelevance}}$ = 2.84, 95%-CI [2.14, 3.55], $BF_{H0H1} = 6.61*10^{-16} \approx 0$), but no main effect for the Item factor was found (b_{ItemTT} = 0.24, 95%-CI [-0.04, 0.52], $BF_{H0H1} = 8.57$). Presumably, the lack of a main effect of the truth table cells on the reason-

relation ratings is due to the fact that the TT vs. TF difference only affects the Positive Relevance condition.

This combination of effects indicates that it remains the case that Positive Relevance contents are evaluated higher than Irrelevance contents, even in spite of the TT/TF manipulation. However, to assess impact of the TT/TF manipulation more closely, Figures 9a and 9b were made. Figure 9a plots the central tendencies of M10, whereas Figure 9b plots the posterior predictions of M10 for the individual response categories:

Figure 9a. Central Tendencies of M10



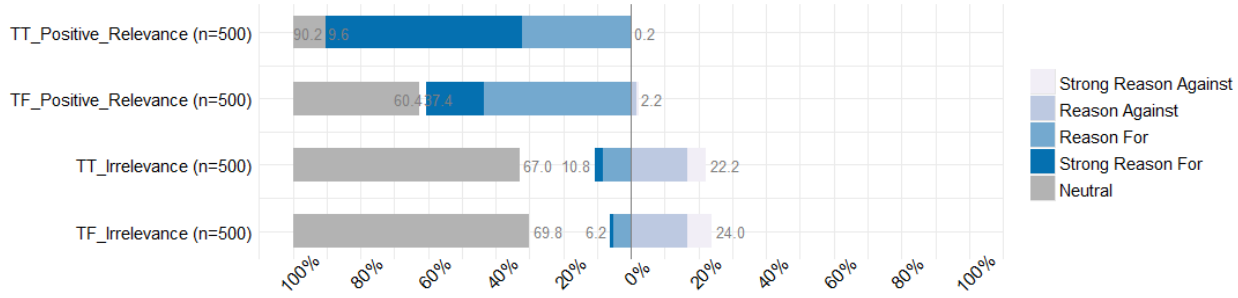
Note. Estimated marginal means for the reason-relation assessments in the different truth table cell x relevance conditions.

As Figure 9a shows, the central tendency remains that the participants agree that the antecedent constitutes a reason for the consequent in the Positive Relevance condition, even when the TF manipulation is added, and that the participants responded that the antecedent is neutral with respect to the consequent on the perceived reason relation scale for the Irrelevance condition.

However, as Figure 9b indicates, for the Positive Relevance condition there is a decrease from ca. 90 % agreeing that the antecedent constitutes a reason or a strong reason for the consequent in the TT cell to ca. 60% of the participants agreeing that it constitutes a reason for or

a strong reason for the consequent in the TF cell. In contrast, for the Irrelevance condition there is almost no change with ca. 67 % agreeing that the antecedent is Neutral with respect to the consequent on the perceived reason relation scale for the TT items, and ca. 70% making the same judgment for the TF items.

Figure 9b. Posterior Predictions of Individual Categories, M10



Note. The perceived strength and direction of the reason relation across truth table cells (TT vs. TF) was measured on a scale from 'strong reason against' (1), 'reason against' (2), 'neutral' (3), 'reason for' (4), and 'strong reason for'.

Yet, as a central tendency it remains the case that the Positive Relevance items are rated one point higher on the reason-relation scale than the Irrelevance items on average, even in the presence of the TF cell.

Discussion

From the analysis we can conclude that while the truth-table cell diminishes the effect of the reason-relation manipulation, it is far from suppressing it. The median response is still to treat A as a reason for C in the Positive Relevance TF condition (in spite of an increase in 'Neutral' response), and for the Irrelevance conditions the truth-table cells did not have any effect.

Hence, alternative accounts of our findings based on the objection that the presence of the truth-table cells undermines the reason-relation manipulation are not supported by the data.

General Discussion

In this paper, we have been concerned with diagnosing whether the reason-relation reading is due to a pragmatic or semantic component of indicative conditionals. In addressing this question, we have empirically contrasted it with well-known linguistic phenomena - namely, conversational implicatures, entailments, presuppositions, and conventional implicatures. In the course of these investigations, we discovered that 1) attempting to cancel a commitment to the reason-relation reading of indicative conditionals is viewed as just as contradictory as cancelling a commitment to an entailment (whereas attempting to cancel a commitment to the reason-relation reading of conjunction is viewed as less contradictory than cancelling a commitment to a scalar implicature), 2) to negate a conditional in wide scope is not in general viewed as equivalent to negating the consequent of a conditional, and 3) the reason-relation reading can become at-issue content not only for conditionals but also for therefore-sentences, which constitute a paradigmatic example of Gricean conventional implicatures.

Can Mental Model Theory explain our Data?

As we pointed out in the introduction, the Mental Models Theory does not treat the relation between the antecedent and consequent of a conditional as a part of its core meaning. However, it postulates "a mechanism of *modulation* that can transfer [the core] meaning into an indefinite number of different sorts of interpretation" (Quelhas, Johnson-Laird, and Juhos, 2010, p. 1717). More specifically, which mental models are and which are not constructed in the process of the interpretation of a conditional can be affected by the semantics, that is, the meanings of its clauses, or by pragmatics, that is, the knowledge about the context of utterance and general world knowledge related to what the conditional is about. For instance, modulation can result in blocking the construction of a model corresponding to the possibility of ¬A and C which yields the bi-conditional interpretation. To give an example, in a context in which a TV is off, an

interpretation of the conditional "If Mark presses the on switch, his TV will be on" might consist of the following models:

switch on TV on

switch off TV off

The possibility that Mark doesn't press the on switch but the TV is on is here excluded by the background knowledge (see Quelhas et al. 2010, p. 1720, for a list of possible interpretations of the conditional). Semantic modulation can also prevent us from accepting conditionals such as:

(4) If God exists, then atheism is correct.

Here, the meaning of "atheism is correct" entails that God does not exist, the mental model consisting of the antecedent and consequent of this conditional is not possible, and hence the conditional is false (Johnson-Laird et al. 2015, p. 206, Quelhas, Rasga, and Johnson-Laird, 2017, p. 24). Note, however, that here the mechanism of modulation takes the presence of the analytic relationship between the clauses of a conditional as its input, rather than an output. In other words, the presence of an analytic connection between the words used in the antecedent and consequent is the reason why the construction of certain mental models is blocked, thus blocking the construction of mental models (comparable to deleting rows in the graphical representation of the explicit model) does not explain why a broader range of conditionals expresses the existence of a connection.

More importantly, the type of a connection conveyed by the conditionals we investigated goes beyond such analytic relationships. Granted, the advocates of the Mental Models Theory propose that the mechanism of modulation can also add information to the model of a sentence. In particular, it can add relations between the clauses of a conditional (see, e.g., Johnson-Laird &

Byrne, 2002, p. 651, Quelhas et al., 2010, pp. 1728-9, Khemlani et al., 2018, pp. 12-13). For instance, Quelhas et al. (2010, p. 1728) observe that:

> appropriate contents should introduce a temporal relation between antecedent and consequent events—for example:
>
> (5) If Lisa received some money, then she paid Frederico.
>
> Individuals know that payment can be made only if a payer has money, and so modulation should yield an interpretation of the conditional in which if Lisa received some money then she did so before she paid Frederico.

Given the above description, the mechanism of modulation involved in this type of case can be construed as some kind of a pragmatic inference: the form of a conditional plus general background knowledge relevant for the interpretation of its antecedent and consequent allow people to *infer* the temporal order (or, in other cases, spatial relations, causal dependencies, and other possible relationships) of the events the antecedent and consequent are about. In other words, unless we deal with the analytical relationships between words that occur in the antecedents and consequents, what is responsible for the variety of different interpretations of conditionals, on the Mental Models Theory, are pragmatic processes. Consequently, while the mechanism of modulation tells us where the perceived connections between antecedents and consequents might be coming from, the Mental Models Theory has no resources to explain why, among other things, participants judge such connections not to be cancellable without giving rise to a contradiction, nor why participants tend to find conditionals whose antecedents and consequents are not connected to be somehow defective.

Finally, we would like to note that the mechanism underlying modulation described in Johnson-Laird & Byrne (2002) and Khemlani et al. (2018) relies on deleting or adding rows of the truth table, if there are salient pragmatic factors, or lexical content in the clauses of the conditional, which exclude these truth-table cells (like in the examples discussed above). Yet part of the data, which the conventional implicature hypothesis is introduced to account for, is the strong dissociation in terms of the influence of relevance on probability/acceptability evaluations and the lack of influence of relevance on truth evaluations (Skovgaard-Olsen et al., 2017). Since this finding shows that the impact of relevance on conditionals is mostly found for types of cognitive assessments other than truth evaluations, an explanation which posits that truth-table cells are deleted or added in missing-link conditionals is unlikely to account for the complex data pattern. Consequently, semantic and pragmatic modulation of the type described above cannot explain the participants' reaction across experiments that conditionals whose antecedents and consequents are not connected are somehow defective.

Repercussions for Argumentation with Conditionals

In and of themselves, our findings have interesting repercussions for argumentation with conditionals. What they suggest is that uttering regular[13] indicative conditionals commits speakers to there being a reason relation between its antecedent and consequent, which speakers cannot escape from without retracting the original utterance, if they are to avoid leaving the impression that they are contradicting themselves. Moreover, when the interlocutor negates a speaker's conditional assertion "If A, then C", then the interlocutor need not be taken as

[13] The qualification is meant to set aside the problematic case of biscuit conditionals (e.g. 'if you are hungry, there are biscuits on the sideboard') for present purposes. For further discussion see Biezma and Goebel (*in review*).

committing to "If A, then ¬C". Finally, the results indicate that there are situations where the reason relations expressed by both conditionals and 'therefore' sentences may become content-at-issue, and indeed that there is a stronger tendency to take these reason relations as content-at-issue than the truth values of their constituents. This in turn suggests that the interlocutor need not interrupt the natural flow of the conversation by expressions like "Hey, wait a minute!" to challenge a reason relation expressed by the speaker through indicative conditionals or 'therefore' sentences. Rather, the interlocutor can challenge the speaker's reason-relation commitments directly and treat these commitments as the main point of the assertion.

To illustrate how all these phenomena may shape argumentative discourse, consider the example of a European conference on global warming in the winter of 2010-2011, which was notorious for featuring an unusually cold November in Central Europe, in some cases setting records for low temperatures. Or alternatively: think of the winter of 2017-2018 in the USA on the East coast. Suppose a politician representing climate change skepticism utters 'if global warming is real, then winter will be warmer than we are used to'. His interlocutors in turn might want to negate this utterance, without thereby committing themselves to the claim that 'if global warming is real, then winter will be colder than we are used to', which they likewise reject. According to the Negation Principle such a discourse move would, however, have been incoherent, whereas the present considerations and our results suggest that it is not.

Suppose further that a consensus forms at the conference that claims about global warming are concerned with global, long-term climate trends, which are unaffected by temporary, regional weather events. In that case, the politician from above might attempt to back-pedal and eschew a commitment to a reason relation between global warming and local weather events by performing a cancellation speech act. However, the results from Experiment 1

indicate that he or she would likely be viewed as contradicting himself/herself. Instead the politician would have to admit an error by retracting his/her earlier statement.

Finally, the results from Experiment 2 suggest that reason-relation commitments like the one above are not the sort of thing that tacitly get introduced into the common ground through indicative conditionals and therefore-sentences without themselves being the subject of direct rejections and affirmations. When it comes to controversial topics like climate change, which reason relations to accept is itself disputed territory and can become the main point of assertion. It would be a natural continuation of the discourse to make disputed reason relations the content-at-issue by targeting them with direct denial and affirmations. The interlocutors need not, in other words, interrupt the natural flow of conversation to challenge reason-relation commitments and we need not conceptualize reason-relations as uncontroversial assumptions that are automatically accommodated into the common ground.

Conventional Implicatures and At-Issue Content

As explained in Koev (2018), there are various notions of at-issue content discussed in the literature and it is, presumably, unclear how they are to be unified. One important property highlighted in Potts (2005, 2007) is that at-issue content constitutes direct proposals to update the common ground of mutually shared assumptions by the interlocutors. On this view, at-issue content is negotiable and open to direct agreement or disagreement by the addressee, and a key diagnostic is whether discourses are acceptable where the interlocutor provides a direct response ("Yes, .../No, ...") to the target content. In contrast, content not-at-issue is thought of as grammatically encoded content that is directly imposed on the common ground without negotiation.

A common assumption is that discourses are structured around questions under discussion. A second important property of at-issue content is that it provides potential answers to the question under discussion (Koev, 2018).

Following Bach (1999), Potts (2005, 2007) used the first notion of at-issue content in his treatment of appositives, which he treats as a paradigmatic instance of conventional implicatures. On this view, conventional implicatures are "secondary entailments that cooperative speakers rarely use to express controversial propositions or carry the main themes of a discourse" (Potts, 2007: 476). For instance, in 'Mozart, the famous composer, used to live here', the assumption that Mozart was a famous composer is presented as a shared assumption not really up for discussion. It is grammatically marked as not the central focus of the assertion. Yet, its content is conventionally part of the meaning of the sentence, and not produced by, say, a conversational implicature.

In continuation of this line of work, Experiment 3 set out to probe whether the reason-relation reading of conditionals is content-at-issue in an attempt to determine whether the reason-relation reading is a conventional implicature. The results showed that the participants clearly treated the reason-relation reading as content-at-issue for conditionals. Interestingly, we found the same pattern of results with respect to therefore-sentences, which Grice (1989) treated as a paradigmatic instance of conventional implicatures.

Our results thus stand in tension with the newer literature on conventional implicatures which treats not-at-issue content as a diagnostic feature. However, as shown in Salmon (2011), it turns out that Pott's (2005, 2007) notion of a conventional implicature differs subtly from Grice's (1989) both in terms of its properties and central instances. Gricean conventional implicatures are non-truth conditional, non-cancellable, not calculable based on the Gricean maxims of

conversation, and detachable (e.g. substituting 'and' with 'but' in the following sentence in the same context of utterance will lose the implication of a contrast between being poor and honest: "She was poor but honest"). In contrast, Pott's conventional implicatures are also speaker-oriented in that the speaker incurs a commitment to them even when making indirect speech reports (e.g. in "John said that Ames, the former spy, is now behind bars", the speaker is also committed to Ames being a former spy). Gricean conventional implicatures lack this property (Salmon, 2011). This fits with the idea of reason relation as Gricean conventional implicatures, because in neither of the two following examples is it the case that the speaker incurs a commitment to a reason relation by making an indirect speech report: “The politician said that if global warming is real, then winter will be warmer than we are used to”, “The politician said that the winter is surprisingly cold therefore global warming is bogus”.

A further difference between the two notions of conventional implicatures may be that Potts's notion (which centers around appositives and non-restrictive relative clauses) differs from Grice's (which centers around utterance modifiers like 'therefore', 'but', and 'even')[14] exactly with respect to the at-issue status of the content.

A first indicator that something is amiss is that in a textbook example like "she is poor but honest", it is far from the case that something uncontroversial, which is not up for negotiation, is expressed by the implied contrast between poverty and honesty. Furthermore, the skeptic’s assertion of 'if global warming is real, then winter will be warmer than we are used to' is used to express a highly controversial claim. There is no reason why his interlocutors should treat it as part of the common ground which is not up for negotiation. The same would apply if he had

[14] If the argument in Salmon (2011) goes through, then the indirectness of the evidence possessed signalled by the epistemic modal 'must' can be added to the list.

formulated his statement as 'The winter of 2010-2011 is unusually cold; *therefore* climate change is bogus'. Moreover, it holds for both these statements that the speaker clearly intends them as partial answers to the question under discussion of the conference. Indeed, we would expect these statements to elicit a discussion about whether the politician can really use local weather phenomena like the winter of 2010-2011 as decisive evidence against climate change without interrupting the natural flow of the discourse at the conference.

The results from Experiment 3 indicate that the participants do treat the reason-relation reading of conditionals and 'therefore' sentences as content-at-issue. Perhaps one reason is the worry that Pierre, the foreign language learner, might advance controversial, nonsensical reason relations unless he is directly confronted. Moreover, the presence of justifications in the answer options that explicitly target the presence or absence of reason relations in Pierre's utterance may also have contributed to making the reason relations content-at-issue.

What the considerations above suggest is that lack of at-issue content may not be a good characterization of (Gricean) conventional implicatures to begin with. For argumentative discourse we need some way of coordinating which reason relations to accept. When opposing views clash a central part of the dispute is which arguments to accept. If argumentative discourses always came down to which factual premises to accept, then they could be resolved by identifying the most reliable source of evidence on the topic and simply accepting its verdicts.

The picture that emerges out of these considerations is that the reason-relation reading of indicative conditionals is a conventional implicature, which is tapped into through probability and acceptability evaluations. Yet at times reason relations may become content at-issue in the sense of addressing the question under discussion and constituting direct proposals to update the common ground of mutually shared assumptions by the interlocutors. Since, however, (non-

deductive) reason relations constitute probabilistic constraints (i.e. $P(C|A) - P(C|\neg A) > 0$), what this requires is that we go beyond Stalnaker (2016) in thinking of the common ground in terms of a set of propositions by enriching it with probabilistic structure. But this is something that we anyway have ample reason for doing when modeling epistemic and doxastic content (Yalcin, 2012).

Conclusion

In relation to the diagnostic problem with which we started, our investigations permit us to draw the following conclusions.

The Relevance Effect reported in Skovgaard-Olsen et al. (2016a) is probably not due to the influence of a conversational implicature. From Experiment 1 no support could be derived for the hypothesis that the reason-relation reading of indicative conditionals is generated by the presence of a conversational implicature. However, support could be obtained that the reason-relation reading of conjunctions could be the result of a conversational implicature. Experiment 1 thus also contributes to drawing a dividing line between the content of logical operators. In discussions, the argument is often put forward that the Relevance Effect on conditionals cannot be taken to reveal something about the semantic content of conditionals, because conjunctions also have a reason-relation reading and presumably we would not want to make it part of the semantic content of conjunctions. The results of Experiment 1 directly undercut any such argument by showing the different status that the reason-relation reading of indicative conditionals and conjunctions have with respect to conversational implicatures.

In the discussion of these results, we considered, however, alternative interpretations of our results based on varying strengths of conversational implicatures, and the possibility of other ways of phrasing the cancellation task. In response, it was pointed out that these alternative

hypotheses are confronted with the burden of explaining why our way of posing the task worked so well with our comparison cases such as the conjunctions, if the results with respect to conditionals are not taken at face value.

A further conclusion of this paper is that the Relevance Effect on conditionals is probably not due to the presence of a presupposition failure of the irrelevance items. One of the most characteristic properties of presuppositions is their ability to project when embedded under logical operators, like negation. However, Experiment 2 could not find support for the hypothesis that the reason-relation reading of indicative conditionals projects when embedded under negations. And perhaps even more damaging to the presupposition failure hypothesis is the finding that while extreme individual differences could be found in the probability assignment to control items with presupposition failures, these individual differences were not reflected in the participants' probability assignment to missing-link conditionals. In our discussion of these results, we again considered, but rejected, an alternative interpretation of our results based on local accommodation.

Experiment 2 moreover yielded a further finding of interest in its own right. According to a well-known negation principle employed in various systems of conditional logic, wide scope negation equals narrow scope negation (Arlo-Costa, 2007). This principle has centrally figured in Suppositional Theory of conditionals' account of compound conditionals involving negations (Edgington 1997, 2000, 2006; Woods, 1997, ch. 6; Kölbel, 2000; and Bennett, 2003, ch. 7). Indeed, in Handley et al. (2006) this principle is even treated as a litmus test for the suppositional conditional. However, the data clearly show that while this negation principle can be maintained for the Positive Relevance condition, it is systematically violated for the Irrelevance condition.

Hence, Experiment 2 provides an occasion to reevaluate how indicative conditionals interact with the negation operator.

Turning to Experiment 3, the reason-relation reading of indicative conditionals was found to be treated as more at-issue than the truth functionality of the clauses. However, this latter finding was also found with 'therefore' sentences. And, indeed, if we take the present results together with those in Skovgaard-Olsen et al. (2017), then we see that indicative conditionals and therefore-sentences behave in a similar way in a range of cognitive assessments which have a bearing on whether the Relevance Effect is a conventional implicature. To be sure, we did not find for either indicative conditionals or for therefore-sentences that the reason-relation reading was not-at-issue content. But it was found in Skovgaard-Olsen et al. (2017) for both types of sentences that the reason-relation reading shows strong fingerprints in tasks where probability or acceptability is asked; however, when the participants are asked to fill out truth-tables they almost entirely ignore the reason-relation reading.

And so, the argument for the (Gricean) conventional implicature interpretation of the Relevance Effect now stands supported by a) the negative results for the relevance effect being a conversational implicature, b) the negative results for the relevance effect being a presupposition failure, c) the finding of the dissociation with relevance strongly affecting probability and acceptability but hardly affecting truth value assignments in the truth table task, d) the fact that indicative conditionals behave remarkably like 'therefore' sentences when probed for their truth, acceptability, probability, or when compared on their at-issue content.

In our discussion, we considered further alternative hypotheses with respect to a) and b) and expressed our doubts. But we welcome future empirical studies that may challenge these

conclusions through variations of experimental tasks and attempts to provide unifying explanations of the complex data pattern that is emerging.

In the meantime we conclude that our data suggest that the Relevance Effect is not a conversational implicature and is not due to presupposition failure. The best candidate, instead, is most likely a conventional implicature. These findings suggest a new direction for the debate on whether relevance is part of the semantics or pragmatics of the conditional. A final judgment will rest on the definition of semantics and pragmatics, and on how conventional implicatures are categorized according to that definition. It may be, however, that conventional implicatures - and, hence, the reason relation reading of conditionals - remain an intermediate, irreducible layer of meaning between semantics and pragmatics.

References

Arlo-Costa, Horacio (2007). The Logic of Conditionals. *The Stanford Encyclopedia of Philosophy* (spring 2016 Edition), Edward N. Zalta (ed.). URL = <http://plato.stanford.edu/archives/fall2016/entries/logic-conditionals/>.

Baayen, R. H., Davidson, D. J., and Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, *59*, 340-412.

Baratgin, J., Douven, I., Evans, J. St. B. T., Oaksford, M., Over, D.E., & Politzer, G. (2015). The new paradigm and mental models. *Trends in Cognitive Sciences*, 19, 547–548.

Baratgin, J., Politzer, G, Over, D. E., and Takahashi, T. (2018), The Psychology of Uncertainty and Three-Valued Truth Tables, *Frontiers in Psychology, 9:1479.*

Batchelder, W. H., & Riefer, D. M. (1999). Theoretical and empirical review of multinomial process tree modeling. *Psychonomic Bulletin & Review, 6*, 57–86.

Beaver, D. I. and Geurts, B. (2014). Presupposition. In E. N. Zalta (Eds.), *The Stanford Encyclopedia of Philosophy*, URL = <http://plato.stanford.edu/archives/win2014/entries/presupposition/>.

Bennett, J. (2003). *A Philosophical Guide to Conditionals*. Oxford: Oxford University Press

Biezma, M. (2014). The grammar of discourse: The case of *then*. In T. Snider et al. (eds.), *Proceedings of SALT 24* (pp. 373-394). Cornell U: LSA and CLC Publications.

Biezma, M. and Goebel, A. (*in review*). Being pragmatic about biscuits. Retrieved from https://mariabiezma.com/biezma-some-recent-work/

Birner, B. J. (2014). *Introduction to Pragmatics*. Malden, MA: Wiley-Blackwell.

Blome-Tillmann, M. (2013). Conversational Implicatures (and How to Spot Them). *Philosophy Compass*, 8/2, 170-85.

Bürkner, P. (2017). brms: An R Package for Bayesian Multilevel Models Using Stan. *Journal of Statistical Software*, *80*, 1-28.

Bürkner, P., & Vuorre, M. (2018, June 23). Ordinal Regression Models in Psychology: A Tutorial. https://doi.org/10.31234/osf.io/x8swp

Carston, R. (1993). Conjunction, explanation and relevance. *Lingua*, 90, 27-48.

Carston, R. (2002). *Thoughts and Utterances*. Oxford: Blackwell Publishers.

Chierchia, G. and McConnell-Ginet, S. (1990). *Meaning and Grammar*. Cambridge, MA: the MIT Press.

Dancygier, B. and Sweetser, E. (2005). *Mental Spaces in Grammar: Conditional Constructions*. Cambridge University Press, Cambridge.

Davis, W. (2014). "Implicature", *The Stanford Encyclopedia of Philosophy* (Fall 2014). In: Zalta, E. N. (Ed.). Retrieved from https://plato.stanford.edu/archives/fall2014/entries/implicature/

Declerck, R. and Reed, S. (2001). *Conditionals: A Comprehensive Empirical Analysis*. Mouton de Gruyter: Berlin/New York.

Douven, I. (2015). *The Epistemology of Indicative Conditionals. Formal and Empirical Approaches*. Cambridge: Cambridge University Press.

--- (2017). How to account for the oddness of missing-link conditionals. *Synthese,* 194 (5), 1541-1554.

Douven, I. and Verbrugge, S. (2010). The Adams family. *Cognition*, *117*, 302–318.

Edgington, D. (1995). On conditionals. *Mind*, *104*(414), 235-329.

--- (1997). Commentary. In: Woods, M. *Conditionals*. Oxford: Oxford University Press, 95-138.

--- (2000). General Conditional Statements: A Response to Kölbel. *Mind*, 109 (433), 109-116.

--- (2006). Conditionals. In: Zalta, E. N. (ed.),*The Stanford Encyclopedia of Philosophy*. (Winter 2008 Edition).

URL = <http://plato.stanford.edu/archives/win2008/entries/conditionals/>.

Egré, P., & Politzer, G. (2013). On the negation of indicative conditionals. In M. Aloni, M. Franke, and F.Roelofson (Eds.), *Proceedings of the Amsterdam Colloquium* (pp. 10-18).

Espino, O., and Byrne, R.M.J. (2012). It's not the case that if you understand a conditional you know how to negate it. *Journal of Cognitive Psychology, 24*(3), 329-334.

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.

Fugard, J. B., Pfeifer, N., Mayerhofer, B., & Kleiter, G. (2011). How people interpret conditionals: Shifts towards the conditional event. *Journal of Experimental Psychology:Learning, Memory,and Cognition, 37*, 635–648.

Gazdar, G. (1979). *Pragmatics: Implicature, Presupposition and Logical Form*. New York: Academic Press.

Grice, H.P. (1989). *Studies in the Way of Words*. Cambridge, MA: 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, and Cognition, 32*(3), 559-569.

Heim, I. (1983). On the projection problem for presuppositions. In Barlow, M. and Flickinger, D. and Westcoat, M. (eds.), *Second Annual West Coast Conference on Formal Linguistics*. Stanford University, 114–126.

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

Jeffrey, R. C. (1991). Matter of fact conditionals. *Aristotelian Society Supplementary Volume*, *65*, 161–183.

Johnson-Laird, P. N. and Byrne, R. M. J. (1991). *Deduction*. Hilsdale, NJ: Lawrence Erlbaum Associates Inc.

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

Johnson-Laird, P. N. and Khemlani, S. S. (2014). Toward a Unified Theory of Reasoning. *Psychology of Learning and Motivation*, *59*, 1-42.

Johnson-Laird, P. N., Khemlani, S. S., and Goodwin, G. P. (2015). Logic, probability, and human reasoning. *Trends in Cognitive Science*, *19*(4), 201-214.

Kadmon, N. (2001). *Formal Pragmatics*. Malden, MA: Blackwell Publishers.

Karttunen, L. (1973). Presuppositions of Compound Sentences. *Linguistic Inquiry*, 4, 167–193.

Khemlani, S., Byrne, R. M. J., and Johnson-Laird, P. N. (2018). Facts and Possibilities: A Model-Based Theory of Sentential Reasoning. *Cognitive Science*, *42*(6),1-18.

Khemlani, S., Orenes, I., & Johnson-Laird, P.N. (2012). Negation: A theory of its meaning, representation, and use. *Journal of Cognitive Psychology, 24*, 541-559.

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

Koev, T. (2018). Notions of At-issueness. *Language and Linguistics Compass***.** doi 10.1111/lnc3.12306. (*in press*)

Kölbel, M (2000). Edgington on Compounds of Conditionals. *Mind*, 109(433), 97-108.

Kratzer, A. (1986). Conditionals. *Chicago Linguistics Society*, *22*(2), 1–15.

Kruschke, J. (2014). *Doing Bayesian data analysis: A tutorial with R, JAGS, and Stan*. Academic Press.

Khemlani, S. S., Byrne, R. M. and Johnson-Laird, P. N. (2018), Facts and Possibilities: A Model-Based Theory of Sentential Reasoning. Cognitive Science, 42: 1887-1924. doi:10.1111/cogs.12634

Krzyżanowska, K. (2015). *Between "If" and "Then": Towards an empirically informed philosophy of conditionals*. PhD dissertation, Groningen University. Retrieved from http://karolinakrzyzanowska.com/pdfs/krzyzanowska-phd-final.pdf

Krzyżanowska, K, Collins, P. J. and Hahn, U. (2017a). Between a conditional's antecedent and its consequent: Discourse coherence vs. probabilistic relevance. *Cognition*, *164*, 199-205.

--- (2017b). The Puzzle of Conditionals with True Clauses: Against the Gricean Account. In M. Gunzelmann, A. Howes, T. Tenbrink, & E. Davelaar (Ed.), *Proceedings of the 39th Annual Meeting of the Cognitive Science Society* (pp. 2476-2481). London, UK: Cognitive Science Society.

--- (2018). True clauses and false connections. Unpublished manuscript.

Krzyżanowska, K., Wenmackers, S., and Douven, I. (2014). Rethinking Gibbard's Riverboat Argument. *Studia Logica 102*(4), 771-792.

Lassiter, D. (2012). Presuppositions, provisos, and probability. *Semantics & Pragmatics*, *5*, 1-37.

Lewis (1973). *Counterfactuals*. Oxford: Basil Blackwell.

Lewis, D. 1976. Probabilities of conditionals and conditional probabilities. *Philosophical Review*, *85*, 297–315.

McCawley, J. D. (1993). *Everything that Linguists have Always Wanted to Know about Logic, but were ashamed to ask*. Chicago: The University of Chicago Press. (Second edition)

McElreath, R. (2016). *Statistical Rethinking.* Boca Raton, FL: CRC Press.

Moss, S. (2015). On the Semantics and Pragmatics of Epistemic Vocabulary. *Semantics and Pragmatics*, *8*(5), 1-81.

Oaksford, M. & Chater, N. (2007). *Bayesian Rationality: The Probabilistic Approach to Human Reasoning*. Oxford: Oxford University Press.

Oaksford, M. and Over, D. and Cruz, N. (2018). Paradigms, possibilities and probabilities: Comment on Hinterecker et al. (2016). *Journal of Experimental Psychology: Learning, Memory, & Cognition*. (In Press)

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*, 680–693.

Olsen, N. S. (2014). *Making Ranking Theory Useful for Psychology of Reasoning*. PhD dissertation, University of Konstanz.
URL = http://kops.uni-konstanz.de/handle/123456789/29353.

Ospina, R. and Ferrari, S. L. P. (2012). A general class of zero-or-one inflated beta regression models. *Computational Statistics and Data Analysis*, *56*, 1609-1623.

Over, D. E. (2007). New paradigm psychology of reasoning. *Thinking & Reasoning*, *4*, 431-438.

Over, D. E., Hadjichristidis, C., Evans, J. S. B. T., Handley, S. J., & Sloman, S. A. (2007). The probability of causal conditionals. *Cognitive Psychology*, 54(1), 62–97. http://dx.doi.org/10.1016/j.cogpsych.2006.05.002.

Over, D. E., & Cruz, N. (2018). Probabilistic accounts of conditional reasoning. In Linden J. Ball and Valerie A. Thompson (Eds.), International handbook of thinking and reasoning (pp. 434-450). Hove, UK: Psychology Press.

Politzer, G., Over, D. E., & Baratgin, J. (2010). Betting on conditionals. *Thinking and Reasoning, 16*, 172–197.

Potts, C. (2005). *The Logic of Conventional Implicatures*. Oxford: Oxford University Press.

--- (2007). Conventional implicatures, a distinguished class of meanings. In Gillian Ramchand and Charles Reiss (Eds.). *The Oxford Handbook of Linguistic Interfaces* (pp. 475-501). Oxford: Oxford University Press.

--- (2008). Wait a minute! What kind of discourse strategy is this? (Annotated data set) Retrieved from http://christopherpotts.net/ling/data/waitaminute/.

--- (2015). Presupposition and implicature. In Shalom Lappin and Chris Fox (Eds.), *The Handbook of Contemporary Semantic Theory* (pp. 168-202), 2nd edn,. Oxford: Wiley-Blackwell.

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.

R Core Team (2015). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org/.

Ramsey, F. P. (1929/1990). General propositions and causality. In Mellor, D. H. (Eds.), *Philosophical Papers* (pp. 145-163). Cambridge University Press, Cambridge.

Recanati, F. (2011). *Truth-Conditional Pragmatics*. Oxford: Oxford University Press.

Russell, B. (1905). On Denoting. *Mind*, 14, 479–493.

Salmon, W. (2011). Conventional implicature, presupposition, and the meaning of *must*. *Journal of Pragmatics*, *43*(14), 3416-3430.

Skovgaard-Olsen, N. (2016). Motivating the Relevance Approach to Conditionals, *Mind & Language*, **31** (5), 555-579.

Skovgaard-Olsen, N. (*in review*). The (Lack of) Relevance of 'Then' and the Dialogical Entailment Task.

Skovgaard-Olsen, N., Kellen, D., Krahl, H., and Klauer, K. C. (2017). Relevance differently affects the truth, acceptability, and probability evaluations of 'and', 'but', 'therefore', and 'if then'. *Thinking & Reasoning*, 23 (4), 449-482.

Skovgaard-Olsen, N., Singmann, H., and Klauer, K. C. (2016a). The relevance effect and conditionals. *Cognition*, 150, 26-36. doi:10.1016/j.cognition.2015.12.017

Skovgaard-Olsen, N., Singmann, H., and Klauer, K. C. (2016b). Relevance and Reason Relations. *Cognitive Science*, 41(S5), 1202-1215.

Spohn, W. (2013). A ranking-theoretic approach to conditionals. *Cognitive Science*, *37*, 1074–1106.

Stalnaker, R. (1968) "A Theory of Conditionals," *Studies in Logical Theory, American Philosophical Quarterly*, Monograph: 2, 98-112.

Stalnaker, R. (2016). *Context*. Oxford: Oxford University Press.

Strawson, P. F. (1950). On referring. *Mind*, 59, 320–44. Verbrugge, S., Dieussaert, K., Schaeken, W., Smessaert, H., and Belle, W. V. (2007). Pronounced inferences: A study on inferential conditionals. *Thinking & Reasoning, 13*(2):105–133.

Tonhauser, J. (2012). Diagnosing (not-)at-issue content. In *Proceedings of Semantics of Under-represented Languages of the Americas (SULA) 6* (pp. 239-254). UMass, Amherst: GLSA.

van Tiel, B., van Miltenburg, E., Zevakhina, N., & Geurts, B. (2016). Scalar diversity. *Journal of Semantics, 33*(1), 107-135.

von Fintel, K. (2004). Would you believe it? The king of France is back! Presuppositions and truth-value intuitions. In: Reimer, M. and Bezuidenhout, A. (eds.), *Descriptions and Beyond*. Oxford: Oxford University Press, 269–296.

von Fintel, K. (2008). What is presupposition accommodation, again? *Philosophical Perspectives*, *22*(1), 137-170.

Woods, M. (1997). *Conditionals*. Oxford: Oxford University Press.

Xue, J. and Onea, E. (2011). Correlation between presupposition projection and at-issueness: An empirical study. *Proceedings of the ESSLLI 2011 Workshop on Projective Meaning*, Ljubljana, Slovenia.

Yalcin, S. (2012). Context Probabilism. In Aloni, M., Kimmelman, V., Roelofsen, F., Sassoon, G., Schulz, K., and Westera, M. (Eds.), *Proceedings of the 18th Amsterdam Colloquium* (pp. 12-21).