

# Conditionals and Inferential Connections: Toward a New Semantics\*

Igor Douven<sup>†</sup>

SND / CNRS / Sorbonne University

Shira Elqayam<sup>†</sup>

School of Applied Social Sciences, De Montfort University

Henrik Singmann<sup>†</sup>

Department of Psychology, University of Warwick

Janneke van Wijnbergen-Huitink

Faculty of Philosophy, University of Nijmegen

accepted for publication in: *Thinking and Reasoning*

## Abstract

In previous published research (“Conditionals and Inferential Connections: A Hypothetical Inferential Theory,” *Cognitive Psychology*, 2018), we investigated experimentally what role the presence and strength of an inferential connection between a conditional’s antecedent and consequent plays in how people process that conditional. Our analysis showed the strength of that connection to be strongly predictive of whether participants evaluated the conditional as true, false, or neither true nor false. In this paper, we re-analyze the data from our previous research, now focusing on the semantics of conditionals rather than on how they are processed. Specifically, we use those data to compare the main extant semantics with each other and with inferentialism, a semantics according to which the truth of a conditional requires the presence of an inferential connection between the conditional’s component parts.

Keywords: conditionals; inference; semantics; logic; indeterminacy.

## 1 Introduction

Conditionals are sentences of the form “If  $\phi$ , [then]  $\psi$ ,” with  $\phi$  called “the antecedent” and  $\psi$  “the consequent.” Conditionals are central to our thinking and reasoning, both in everyday life and in scientific contexts. In this paper, we are concerned with the semantics of conditionals, or more precisely, with theories about the truth conditions of conditionals. Decades of debate notwithstanding, we are not even close to an agreement on what those truth conditions may be, or even on whether conditionals have truth conditions to begin with. More exactly, our focus will be on the semantics of so-called indicative conditionals, which are conditionals whose antecedent is in the indicative mood,

---

\*The data that were used for this paper can be downloaded from: <https://osf.io/3uajq/>.

<sup>†</sup>The first three authors contributed equally to the paper and are listed alphabetically.

in contrast to subjunctive conditionals, whose antecedent is in the subjunctive mood. (See Douven, 2016a, chap. 1, for more on this distinction.)

To understand what motivated the present research, consider the following two conditionals:

- (1) a. If Mozart died from food poisoning, then Joe Biden will run for president in 2020.
- b. If the price of oil soars in the next months, then Winston Churchill did not sleep the night before D-Day.

There is something manifestly odd about these conditionals. The most plausible diagnosis of *why* they appear odd is that their antecedent has nothing to do with their consequent. There is no intelligible notion of dependency for which Biden's decision about whether he will run for president in 2020 could be said to depend on whether or not Mozart died from food poisoning, and whether Churchill slept the night before D-Day cannot possibly have depended on how the price of oil will develop in the next months. That some sort of connection between its antecedent and consequent is conveyed by the assertion of a conditional is also strongly suggested by considerations regarding this conditional:

- (2) If the price of oil soars in the next months, then Joe Biden will run for president in 2020.

Reflecting on the process of trying to make sense of (2), we notice that it sets us on the path of looking for possible connections between the oil price's soaring in the next months and Biden's decision *re* seeking the presidency. It is not completely inconceivable that such a connection exists: perhaps the oil price's soaring itself, or something leading to its soaring, will be what eventually makes Biden decide to run. But we will reject (2) if we are unable to think of any credible such connection.

In saying that there is something odd about (1a) and (1b), we left it open whether the defect generating the felt oddness is at the level of semantics or of pragmatics. These conditionals could strike us as odd because they are *false*, or perhaps *neither true nor false*—which would make the defect semantic, having to do with truth and falsity—but alternatively also because they are *infelicitous*, which would make the defect pragmatic, having to do with how we *use* sentences. It might be that by asserting a conditional, we *suggest* (or *implicate*, to use the technical term) the presence of a connection between antecedent and consequent, even though the presence of a connection is not part of the truth conditions of that conditional.

Most semantics for conditionals attribute the oddness of sentences like (1a) and (1b) to a pragmatic inadequacy of such sentences: the assertion of a conditional pragmatically conveys the presence of a connection between antecedent and consequent, which in the case of (1a) and (1b) obviously does not exist. But exactly how the presence of a connection is suggested by the assertion of a conditional—exactly which pragmatic mechanisms might give rise to this suggestion—is a question that proponents of the currently popular semantics have given no more than scant attention to. Reasons to be doubtful about the viability of a pragmatic account are to be found in Douven (2008, 2017), Krzyżanowska, Collins, and Hahn (2017), and Krzyżanowska and Douven (2018). In this paper, we will focus on a semantic explanation of why conditionals like (1a) and (1b) are odd.

In the history of philosophy, a number of attempts have been made to account semantically for why conditionals such as (1a) and (1b) strike us as being odd, by making the existence of a connection between a conditional's antecedent and consequent a necessary condition for its truth. Specifically, various authors have argued that the truth of a conditional requires an *inferential* connection between its antecedent and consequent. See, for instance, the following passage from Mill's *System of logic*:

When we say, If the Koran comes from God, Mahomet is the prophet of God, we do not intend to affirm either that the Koran does come from God, or that Mahomet is really his prophet.

Neither of these simple propositions may be true, and yet the truth of the [conditional] may be indisputable. What is asserted is not the truth of either of the propositions, but the inferriability of the one from the other. (Mill, 1843/1872, p. 91)

Another notable advocate of this view is Ramsey. This may come as a surprise, given that Ramsey's name is usually associated with semantics (like non-propositionalism and Stalnaker's semantics—see the next section) that do *not* make an inferential connection between antecedent and consequent a truth condition for conditionals. Nevertheless, in the following passage he explicitly sides with Mill:

In general we can say with Mill that “If  $p$ , then  $q$ ” means that  $q$  is inferrible from  $p$ , that is, of course, from  $p$  together with certain facts and laws not stated but in some way indicated by the context. (Ramsey, 1929/1990, p. 156)

As historians of logic have shown, the view in effect dates back well beyond Mill, having its roots in the writings of the Stoic philosophers (Kneale & Kneale, 1962, chap. 3; Sandford, 1989, chap. 1).

In the psychological literature, the same view has been defended by Braine (1978; see also Braine and O'Brien, 1991). As Braine (1978, p. 8) puts the core idea, “the logical function of *if-then* is taken to be the same as that of the inference line, that is, *if-then* and the inference line are different notations for indicating the same relation between two propositions.” The standard linguistic treatment of conditionals, due to Lewis (1975) and Kratzer (1986, 1991, 2012), also seems an inferentialist approach, especially when cast in what Lewis (1981) calls a “premise semantics.” In this theory, conditionals involve (often implicit) modal expressions like “must.” Kratzer (2012, p. 9) writes: “The meaning of *must* is related to logical consequence: a proposition is necessary with respect to a premise set if it follows from it.” Then a conditional is true in case its consequent follows from the premise set joined with the antecedent.

That this “inferentialist semantics” has never been widely embraced is mainly due to the fact that it seems very easy to come up with conditionals that are plausibly true even though their consequent does *not* follow from their antecedent. Consider the following conditional about old friends Alice and Bob, who however recently had a fight that many thought had ended their friendship for good:

- (3) If Alice and Bob are jogging together again, they have patched things up.

We have no difficulty imagining circumstances under which we would deem this conditional true, even if those same circumstances would not permit us to rule out entirely that Alice and Bob are jogging together despite *not* having patched things up, say, because they have business matters to discuss which they prefer to do while jogging rather than sitting face to face at a desk. Naturally, though, this would mean that we cannot validly infer the conditional's consequent from its antecedent.

It is important to observe, however, that this possibility rebuts the proposal only if “inference” is taken to mean *deductive* inference (i.e., logical consequence). That is manifestly the correct interpretation as far as the quote from Ramsey is concerned (Douven, 2016a, chap. 2). According to most commentators, it is also the interpretation the Stoics (or at least some Stoic philosophers) had in mind (e.g., Mates, 1953, p. 42; Kneale & Kneale, 1962, p. 161). Braine, and Braine and O'Brien, in the articles cited above, also explicitly commit to a deductive notion of inference. Kratzer (1986, 1991, 2012) does acknowledge that a conditional's antecedent might be related to its consequent via an inference notion other than logical consequence, and neither does the inference-as-deduction interpretation seem right in the case of Mill (Skorupski, 1989, p. 73 f). In any event, the idea that the truth of a conditional requires the existence of an inferential connection is by no means wedded to that interpretation, and may instead refer to a notion of inference that generally encompasses other forms of inference besides deduction (not just in light of inconsistencies).

As a case in point, we mention Krzyżanowska, Wenmackers, and Douven’s (2014) recent proposal of an inferentialist semantics for conditionals in which the connection is postulated to consist of an argument from antecedent to consequent that may involve not only deductive but also abductive and inductive inferential steps, where (roughly) abductive inference is inference based on explanatory considerations and inductive inference is inference based on frequency information (whether specified numerically, as in “ninety percent of the students got an A,” or qualitatively, as in “most of the participants were from Canada”). On their proposal, (3) may be true because, in light of the relevant background knowledge, Alice and Bob having patched things up is part of the best explanation for their jogging together again. (*Why* are they jogging together? Because they are friends again, and they always used to enjoy jogging together.) At a minimum, Krzyżanowska and coauthors show that the general idea of making an inferential connection between antecedent and consequent a requirement for the truth of conditionals is far from hopeless. For related proposals, see Rescher (2007), Spohn (2013, 2015), Nickerson (2015), Skovgaard-Olsen (2016a, 2016b), Vidal and Baratgin (2017), and Markovits, de Chantal, and Brisson (2019).<sup>1</sup>

Throughout, we use “inferentialism” in Krzyżanowska et al.’s sense, and thus as designating the broad idea that the truth of a conditional requires an inferential connection between antecedent and consequent, where this connection may be deductive, inductive, abductive, or consisting of a combination of deductive, inductive, and/or abductive steps. In a previous publication (Douven, Elqayam, Singmann & van Wijnbergen-Huitink, 2018), we proposed a novel theory of conditionals, *Hypothetical Inferential Theory* (HIT), based on two pillars: inferentialism, as the semantic level theory, and Evans’s (2006, 2007) Hypothetical Thinking Theory, a dual-process theory of thinking and decision-making. According to HIT, the way we understand conditionals is by representing their antecedent and consequent as being connected by a relevance-driven, satisficing-bound relationship. HIT adds two psychological principles to inferentialism: the idea that people’s default representation of conditionals is one in which there is an inferential connection between antecedent and consequent (*principle of relevant inference*); and that this connection needs only be strong enough, in the sense suggested by Simon’s (1982) notion of satisficing (*principle of bounded inference*). In the work reported in Douven et al. (2018), we found that the strength of the inferential connection between a conditional’s antecedent and its consequent is a reliable predictor of how people evaluate the conditional’s truth value. Douven et al. explored a psychological theory, and in that context, we did not dwell on the semantic implications of our findings. Specifically, we left an analysis of inferentialism’s adequacy (or lack thereof) *qua* semantics of conditionals to future work. In particular, we did not compare, in light of the data reported in the paper, inferentialism with any of the main extant semantics of conditionals.

The aim of the current paper is to fill this lacuna. In other words, the focus of our current exploration is on the semantics of conditionals. Support for inferentialism as a valid semantic model goes beyond a mere philosophical research question. We submit that it also has high significance for the psychological understanding of conditionals. Inferentialism is the computational-level component of HIT (in the sense suggested by Marr, 1982), that is, it describes the function computed by conditionals. No psychological theory of conditionals is complete without this computational-level analysis. Whether inferentialism is an adequate description of human reasoning is of paramount significance for psychological as well as philosophical theories of conditionals.

In the following, we have a second look at Experiment 1 from Douven et al. (2018), whose design allows such direct comparisons. We report a re-analysis of the data gathered in that experiment,

---

<sup>1</sup>Vidal and Baratgin (2017), and to some extent also Markovits, de Chantal, and Brisson (2019), present experimental results in support of an inferentialist semantics; see also Skovgaard-Olsen et al. (in press). For some apparently countervailing results, see Skovgaard-Olsen et al. (2017).

now explicitly with an eye toward testing inferentialism alongside the main competing semantics of conditionals.<sup>2</sup> We begin, in Section 2, by summarizing the central tenets of the main semantics for conditionals and by stating inferentialism in greater detail than was done in the preceding paragraphs. In Section 3, we describe Experiment 1 from Douven et al. (2018), and in Section 4, we evaluate the bearing of the data from that experiment on the various semantics for conditionals.

## 2 Semantics

### 2.1 The material conditional account

The material conditional account has long been the received view in the semantics of conditionals. Famous proponents include Lewis (1976), Jackson (1979, 1987), and Grice (1989); for a recent defense, see Rieger (2006, 2013). According to this account, the truth conditions of a conditional are those of the corresponding material conditional (see Evans & Over, 2004, and Nickerson, 2015, for discussions in the context of psychological theories of conditionals): “If  $\phi$ ,  $\psi$ ” is false if  $\phi$  is true and  $\psi$  is false, and in all other cases it is true. This account has a number of notable virtues: it is simple, it validates various inferences that are also intuitively valid, and it makes it easy to see how conditionals interact with the rest of the language. But the account also faces serious problems. Most famously, it gives rise to the so-called paradoxes of the material conditional: it validates the intuitively invalid inference of “If  $\phi$ ,  $\psi$ ” from not- $\phi$  (e.g., the inference of “If Paris is the capital of England, Paris is the capital of Turkey” from “Paris is not the capital of England”), as well as the intuitively equally invalid inference of “If  $\phi$ ,  $\psi$ ” from  $\psi$  (e.g., the inference of “If Bill Gates just went bankrupt, he is a billionaire” from “Bill Gates is a billionaire”). While various attempts have been made to explain these problems away—typically along pragmatic lines, as in Grice (1989)—it is fair to say that the account is no longer considered as the received doctrine among theorists working on conditionals (see also Pfeifer & Douven, 2014).

### 2.2 Stalnaker’s semantics

For a while, Stalnaker’s (1968) possible worlds semantics came close to enjoying the status of being the received doctrine, at least in philosophy. According to Stalnaker, a conditional is true (false) if its consequent is true (false) in the closest possible world in which its antecedent is true—provided there is a world in which its antecedent is true; otherwise it is vacuously true. “Closest possible world in which its antecedent is true” is a term of art, but loosely speaking it is the way the world would be if it were minimally different from the actual world so as to make the antecedent true (so, if the antecedent is true in the actual world, then the world closest to the actual world in which the antecedent is true is the actual world itself).

A main selling point of this account is that it seems to capture semantically the intuition underlying the celebrated Ramsey Test for the acceptability of conditionals. In this test, which has great psychological plausibility (Evans & Over, 2004), we determine whether to accept a given conditional by hypothetically adding its antecedent to our stock of beliefs, making minimal changes (if necessary) to preserve consistency, and judging from the resulting (hypothetical) perspective the acceptability of the conditional’s consequent. The closest possible world in which a conditional’s antecedent holds is naturally thought of as a metaphysical correlate of the belief state that results from updating

---

<sup>2</sup>Given that our data consist only of truth evaluations, it is not possible to include in the comparison Jeffrey’s (1991) semantics, according to which “If  $\phi$ ,  $\psi$ ” evaluates to  $\Pr(\psi | \phi)$  in case  $\phi$  is false. This omission is unfortunate, given that Jeffrey’s semantics has been much in the limelight lately; see Over and Baratgin (2017) and references given there.

our current belief state hypothetically on the antecedent. Also, it is easy to verify that Stalnaker's semantics does *not* validate the paradoxes of the material conditional.

Still, this semantics, too, has been severely criticized. For example, many have complained that it fails to validate the Or-to-if principle, which licenses the inference of "If not- $\phi$ ,  $\psi$ " from " $\phi$  or  $\psi$  (or both)." Suppose the disjunction holds because  $\phi$  holds. Nothing follows from that about whether  $\psi$  holds in the closest world in which  $\phi$  is false. This is a drawback if—as critics have argued—the inference from "or" to "if" is intuitively valid. In response to this, Stalnaker (1975) has argued that, although the Or-to-if principle is not valid in his semantics, it can still be said to embody a reasonable inference, in the sense that whenever the disjunction is assertable, so is the conditional. This, according to him, creates the illusion that it is a valid principle. Indeed, applications of the Or-to-if principle are not always intuitively okay, and it has been found that people do not always accept or-to-if inferences (Over, Evans, & Elqayam, 2010; see also Gilio & Over, 2012, and Krzyżanowska, Wenmackers, & Douven, 2014, on when or-to-if inferences are intuitively compelling).

Another general complaint about Stalnaker's semantics is that it fails to validate the so-called Import–Export principle, according to which "If  $\phi$  and  $\psi$ ,  $\chi$ " and "If  $\phi$ , then if  $\psi$ ,  $\chi$ " are interderivable. The principle is widely accepted; at any rate, no one has so far been able to come up with a compelling counterexample to it.<sup>3</sup> But given that the closest  $\phi$ -and- $\psi$ -world may be different from the  $\psi$ -world that is closest to the closest  $\phi$ -world, Import–Export is not valid in Stalnaker's semantics. McGee (1985) has developed a version of possible worlds semantics for conditionals that does validate Import–Export. On the other hand, McGee's semantics fails to validate modus ponens, which many regard as a fatal flaw of the semantics. As McGee (1985:468) notes, however, on his semantics modus ponens may still be a *near*-valid rule of inference, in the sense that, when applied to true premises, the conclusion it yields is *virtually always* true. In his view, this may be all there is to our intuition that modus ponens is a valid rule of inference.

### 2.3 Non-propositionalism

A third semantics for conditionals that has received a fair amount of attention in the literature is non-propositionalism. In this semantics, conditionals do not express propositions; instead they are like requests, commands, promises, and exhortations in that they never have a truth value. According to some advocates of this view, to assert or accept a conditional is to assert or accept its consequent in the event of its antecedent; if the antecedent turns out to be false, nothing has been asserted or accepted (Adams, 1965, 1975). Other advocates hold that conditionals are really truncated arguments, with the conditional's antecedent as premise and the conditional's consequent as conclusion; and arguments can be many things, but not true or false (Mackie, 1973).

It might be wondered how conditionals can play a role in our reasoning if they cannot have a truth value, given that validity is standardly defined in terms of truth preservation. (In standard logic, an inferential principle is said to be valid if and only if the truth of its conclusion is guaranteed by the truth of its premises.) However, non-propositionalists have defined a notion of validity in terms of high-probability preservation and proved that every argument that is classically valid is also valid in this other sense (Adams, 1966, 1975, 1998; Adams & Levine, 1975). To allow this semantics to cover also arguments involving (indicative) conditionals, it was proposed that the probability of a conditional equals the probability of the conditional's consequent given its antecedent. This "Equation," as it is now often called, has since been shown in many experiments to be descriptively adequate. See, among many others, Evans, Handley, and Over (2003), Oaksford and Chater (2003, 2007), Over and

---

<sup>3</sup>This is not to say that it is entirely uncontroversial nor that all semantics other than Stalnaker's do validate it. Most notably, it is invalid in Jeffrey's semantics referenced in note 2; see Gilio and Sanfilippo (2014).

Evans (2003), Evans and Over (2004), Oberauer, Weidenfeld, and Fischer (2007), Over et al. (2007), Douven and Verbrugge (2010, 2013), Pfeifer and Kleiter (2010), Politzer, Over, and Baratgin (2010), Fugard et al. (2011), Over, Douven, and Verbrugge (2013), and van Wijnbergen-Huitink, Elqayam, and Over (2015); see Nickerson (2015) and Over and Cruz (2018) for general discussions of these results, and Skovgaard-Olsen, Singmann, and Klauer (2016) for an important qualification.

Given this proposal, it is easy to show that the semantics at issue does not validate the paradoxes of the material conditional; for instance, assigning a high probability to Paris not being the capital of England does not force one to assign a high conditional probability to Paris being the capital of Turkey on the supposition that Paris is the capital of England. The semantics does not validate the Or-to-if principle either: if it is highly probable that the butler did it then it must be highly probable that either the butler did it or the maid did; but it might still be not highly probable—it might even be highly improbable or simply out of the question—that the maid did it conditional on the butler’s not having done it.

It might be thought that non-propositionalism could be easily disconfirmed in psychological experiments, because people often do apply the terms “true” and “false” to conditionals. The problem is that “true” and “false” could be held to have many non-factual pragmatic uses in ordinary discourse. To say that a conditional, or even a categorical statement, is true could sometimes be merely to endorse it because of subjective probability or utility judgments. People apply “true” even to expressions of pure subjective taste; for instance, they might respond with “true” to “You should have white wine with fish” and to “If you make *coq au vin*, you should use red and not white wine” (Politzer, Over, & Baratgin 2010; see Adams, 1998, on the pragmatic account of truth). In fact, one way to look at non-propositionalism is that, for it, indicative conditionals are closely comparable to deontic conditionals and deontic expressions in general, which can only be “true” in some subjective sense.

However, the arguably most serious problem besetting non-propositionalism is that it has no plausible account of how conditionals combine with other types of sentence in the language, notably, about how they can be constituents of more complex sentences. The problem here is that conjunction, disjunction, and negation are *propositional* operators, taking as their operands only propositions—which conditionals are not, on the present view. Thus, according to non-propositionalism, none of the following sentences should make any sense to us:

- (4) a. Fred will come, and if Anna comes, Jim will come, too.
- b. Either we go to the cinema or, if we don’t like the movies they show, we go to the theater.
- c. If you take your umbrella you will stay dry if it starts raining.
- d. If Susan gets angry if Harry comes home with a B, she will get furious if he comes home with a C.

In practice, these and similar sentences pose not the least interpretational difficulties.

Non-propositionalists have tried to argue that sentences like (4a)–(4d) can all be reduced to sentences that do not involve any embeddings of conditionals. But these attempts have been largely unsuccessful, with conditionals like (4d) appearing particularly recalcitrant to reformulation; see Douven (2011, 2016a, chap. 2, 2016b). Note also that, while non-propositionalists can *endorse* the Import–Export principle—perhaps on the grounds that it sounds *prima facie* plausible—it cannot be said to be *valid* on their account. Specifically, it cannot be claimed that whenever “If  $\phi$  and  $\psi$ ,  $\chi$ ” is highly probable, then so is “If  $\phi$ , then if  $\psi$ ,  $\chi$ .” Assuming the Equation, this would amount to

claiming that whenever  $\Pr(\chi \mid \phi \text{ and } \neg\psi)$  is high, then so is  $\Pr(\text{If } \neg\psi, \chi \mid \phi)$ . But if conditionals do not express propositions then the latter term is undefined.<sup>4,5</sup>

## 2.4 The three-value view

There is a more moderate version of non-propositionalism, which we shall call “the three-value view,” that may be able to overcome the latter problem. In this version, an indicative conditional is made true by a true antecedent and true consequent, made false by a true antecedent and false consequent, and is made neither true nor false when its antecedent is false. In the latter case, the conditional fails to have a truth value and is in some sense “void” (de Finetti, 1995; Edgington, 1995; Bennett, 2003).<sup>6</sup> Like non-propositionalism, the three-value view can also propose that there are non-factual pragmatic uses of “true” (see again Politzer, Over, & Baratgin, 2010, and also Edgington, 2003, on pleonastic uses of “true”). Of course, although conditionals now have a factual truth value when their antecedent is true, they still do not express genuine propositions in the classical sense, in which sentences expressing propositions are always factually true or false.

On the other hand, if we want our logic to apply to natural language, we may need a richer notion of proposition than the classical one anyway. After all, natural language is rife with indeterminacy, quite independently of what we are going to say about conditionals. Vagueness, presupposition failure, reference failure, future contingencies, and paradoxicality (as it is to be found in sentences like the liar sentence, which says of itself that it is false) are, by most philosophers at least, also thought to give rise to indeterminacy. For example, Strawson (1950) famously argued that “The present king of France is bald” is neither true nor false, given that it presupposes something false (viz., that France presently is a monarchy). So, unless we are happy with a logic that applies only to idealized languages (like the language of mathematics, for which modern logic was originally devised), we need a logic that recognizes, and can handle, sentences that are indeterminate (neither true nor false). This means that we will need truth tables that permit truth value gaps, not just for the conditional-forming operator, but also for the conjunction, disjunction, and negation operators.

Thus, if “partial logic” is the right logic for natural language (Blamey, 1986), then, given the three-value view, the conditional-forming operator is not special after all, and there is in principle no difficulty in understanding how conditionals connect with the rest of the language. (In fairness to non-propositionalism, it is to be noted that that position, too, could avail itself of partial logic in order to account for how conditionals combine with other sentences in the language. Naturally, conditionals would still be linguistic oddballs, given that they are *never* true or false: even if conjunctions or disjunctions or negations may, given partial logic, sometimes lack a truth value, often enough they will have one.) This still leaves open the question exactly which partial logic is the correct logic for our language. For instance, is a conjunction with one indeterminate and one false conjunct false or indeterminate? Or, is a disjunction with two indeterminate disjuncts necessarily indeterminate? What if it is of the form “ $\phi$  or not  $\phi$ ”? A priori considerations seem unable to settle such questions,

---

<sup>4</sup>We are assuming classical probability theory here. By extending that theory with a primitive notion of conditional event, Gilio et al. (2017) obtain an account of conditionals in which  $\Pr(\text{If } \neg\psi, \chi \mid \phi)$  is defined, in spite of the fact that the conditional is still not a proposition in the classical sense (see also Gilio & Sanfilippo, 2013, 2014). How much this is going to help non-propositionalists depends on the extent to which Gilio and coauthors, or others, will be able to clarify the metaphysical status of conditional events, which so far they have not given us much guidance on.

<sup>5</sup>Weirich (in press) presents a number of other a priori reasons for rejecting non-propositionalism.

<sup>6</sup>See Baratgin and Politzer (2016) for an argument to the effect that, for de Finetti, non-propositionalism and the three-value are not separate semantics but rather constitute different levels of epistemic analysis of one and the same semantic phenomenon.



Table 1: Different extensions of the three-value view.

$\varphi$	$\psi$	If $\varphi, \psi$		
		de Finetti	Farrell	Cooper
T	T	T	T	T
T	#	#	#	#
T	F	F	F	F
#	T	#	#	T
#	#	#	#	#
#	F	#	F	F
F	T	#	#	#
F	#	#	#	#
F	F	#	#	#

so researchers have started to investigate them by experimental means (e.g., Cruz et al., 2015; Baratgin et al., 2018).

Also, the standard version of the three-value view pertains strictly to conditionals whose antecedent and consequent both have a determinate truth value. But once truth value gaps are admitted in the semantics, we must reckon with the possibility that a conditional’s antecedent or consequent (or both) are indeterminate or uncertain. Again it is not a priori how to extend the three-value view to cover these cases. Table 1 presents the main options that have been discussed in the literature. (In the table, T stands for “true,” F for “false,” and # for “neither true nor false.”)

Validity can be defined in various ways for partial logic. The classical definition in terms of truth preservation looks pretty hopeless. It predicts that “If  $\varphi, \psi$ ” entails the conjunction of  $\varphi$  and  $\psi$ . Non-falsity preservation, by contrast, validates the paradoxes of the material conditional. Because of this, the most popular choice is to require preservation of both truth and non-falsity (McDermott, 1996). This avoids the paradoxes, but does not validate the Or-to-if inference: if  $\varphi$  and  $\psi$  are both true, so is “ $\varphi$  or  $\psi$  (or both),” but “If not- $\varphi, \psi$ ” is indeterminate in that case.

In the three-value view, the probability of a conditional “If  $\varphi, \psi$ ” can be equated with the ratio of the probability that it is true to the probability that it has a (determinate) truth value, which is the conditional probability of  $\psi$  given  $\varphi$  (Rothschild, 2014). Thus the three-value view could alternatively be paired with Adams’ probabilistic notion of validity, according to which high-enough probabilities for the premises guarantee high probability for the conclusion (Adams, 1965, 1975). This makes the same predictions regarding the paradoxes and the Or-to-if inference as does validity defined in terms of truth and non-falsity preservation, but it is not equivalent to it; see Milne (2012) for a comparison.

The standard version of the three-value view, which does not take into account the possibility that the antecedent or consequent of a conditional may be indeterminate, has nothing to say about the Import–Export principle. But it is easy to verify that any extension that evaluates a conditional with an indeterminate consequent as being indeterminate itself—which is the case in all extensions considered in Table 1—does validate the principle. (At least this is so if validity is defined in terms of preservation of some semantic value or values. If it is defined as high-probability preservation, there is still the problem for non-propositionalism raised in note 4.)

Again like non-propositionalism, most followers of the three-value view in fact adopt the probabilistic definition of logical validity in terms of probability preservation. By this definition, a valid argument cannot take us from high probability in the premises to low probability in the conclusion (or, more precisely, to greater uncertainty in the conclusion than there is in the premises; see the earlier references on page 6). The three-value view and this definition of validity have influenced psychologists who are attempting to account for people’s deductive inferences from their uncertain beliefs (Elqayam & Over, 2013; Evans & Over, 2013). For example, they argue that it allows them to explain why people reject the paradoxes of the material conditional and counterintuitive instances of the Or-to-if principle (Over, Evans, & Elqayam, 2010; Pfeifer & Kleiter, 2010; Gilio & Over, 2012).

## 2.5 Inferentialism

The above are the main semantics for conditionals to be found in the literature. While each has its strengths, each also has its weak points. There is one problem that they all share, which in its most general form can be put by saying that they all validate the inferential principle that goes by the name of “Centering” (Cruz et al., 2016). This principle licenses the inference of “If  $\varphi$ ,  $\psi$ ” from the conjunction of  $\varphi$  and  $\psi$ . Because on the material conditional account, Stalnaker’s account, and the three-value view, a conditional with true antecedent and consequent is true, the truth of “ $\varphi$  and  $\psi$ ” entails, on each of these accounts, the truth of both “If  $\varphi$ ,  $\psi$ ” and “If  $\psi$ ,  $\varphi$ .” To see that Centering also holds for non-propositionalism, note that  $\Pr(\varphi \wedge \psi) = \Pr(\varphi) \Pr(\psi | \varphi) \leq \Pr(\psi | \varphi) = \Pr(\text{If } \varphi, \psi)$ , so that whenever the conjunction is highly probable, the conditional is highly probable as well. (This assumes the earlier-mentioned Equation.)

In view of the considerations (concerning the need for a connection between the components of conditionals) we started this paper with, Centering is a pretheoretically implausible principle. It says that for *any*  $\varphi$  and  $\psi$ , however unconnected they may intuitively be, their joint truth (or the high probability of their conjunction) entails the truth (or high probability) of “If  $\varphi$ ,  $\psi$ .” So, for example, if the antecedents and consequents of (1a) and (1b) happen all to be true, then so are these conditionals themselves, their apparent absurdity notwithstanding.

We have already noted above that there is not yet a published pragmatic account of why (1a) and (1b) are absurd, and that we will not try to develop or critically examine one here (for some relevant points, see Douven, 2008, and Krzyżanowska & Douven, 2018). Our focus in this paper will be on the semantic theory of inferentialism, which, we saw, has been defended by Mill, Ramsey, and others. We also saw that the requirement of an inferential connection between antecedent and consequent is best not spelled out strictly in terms of deduction; that would make many conditionals that are pretheoretically true come out false. It was further seen that this lesson is taken to heart in a recent proposal by Krzyżanowska, Wenmackers, and Douven (2014), which does not insist on a *deductive* inferential connection between antecedent and consequent. In line with what is suggested by the earlier quote from Ramsey, and also in line with recent more general work in semantics, this proposal attributes context-relative truth conditions to conditionals. (See MacFarlane, 2012, for a general discussion of context-relative semantics.) Specifically, the core of the proposal is that a conditional is true in a context if and only if there is an argument from its antecedent plus contextually relevant background knowledge to its consequent, where this argument consists of one or more inferential steps, each of which is valid in a deductive, inductive, or abductive sense.<sup>7</sup>

<sup>7</sup>To avoid spurious debate, it is to be noted that linguists and philosophers have long recognized that there are special classes of conditionals—sometimes called “nonconditional conditionals” (Lycan, 2001) or “unconditionals” (Merin, 2007; Spohn, 2013)—which do not require the existence of a connection between their antecedent and consequent. These include Dutchman conditionals (Jackson, 1979, 1987), non-interference conditionals (Bennett, 2003; Burgess, 2004), and

To this core, they add two further clauses, to wit, first, that the consequent ought not to follow (again in the generalized sense that there is an argument consisting of deductively, abductively, or inductively valid steps) from the background knowledge alone unless it also follows from the antecedent alone, and second, that the antecedent ought to be deductively consistent with the background knowledge unless the consequent follows from the antecedent alone. These additional clauses are meant to ensure that the antecedent is not redundant in the derivation of the consequent, and, respectively, that the consequent does not follow trivially from the antecedent plus background knowledge. Without these clauses, one could still have true conditionals without any intuitive connection between antecedent and consequent.

As Krzyżanowska and coauthors admit, there are currently no accounts of abductive and inductive inference that are as well understood and generally accepted as classical logic, our standard account of *deductive* inference. For present purposes, no detailed assumptions need to be made concerning abductive or inductive inference. Indeed, for our purposes Krzyżanowska and coauthors' proposal may be summarized as stating that a conditional is true if and only if there exists a *strong enough* argument leading from its antecedent plus background knowledge to its consequent. What counts as strong enough may be different for different people and is a vague matter. Then again, even people sharing the same background knowledge may differ about the truth value of a conditional; and it is sometimes vague whether a conditional is true or not (Edgington, 1992, 1997).

Krzyżanowska and coauthors show that their inferentialist semantics has a number of clear virtues. Not only has it no difficulty in blocking the paradoxes of the material conditional—neither the truth of the consequent of a conditional nor the falsity of its antecedent ensures the existence of an inferential link between its antecedent and consequent—it also gets the Or-to-if principle right precisely in the kind of cases in which it is intuitively right (see Krzyżanowska, Wenmackers, & Douven, 2014, for the details). These authors further argue that inferentialism provides a solution to a problem case presented in Gibbard (1981), which Gibbard and others had taken to be strong evidence for non-propositionalism.

Krzyżanowska and coauthors note that, because abduction and induction are ampliative forms of inference—meaning that they do not *guarantee* the truth of a conclusion on the basis of the truth of the premises—modus ponens is, on their account, no longer strictly valid: there is the possibility that a true conditional has a true antecedent and a false consequent, so that applying modus ponens to it and its antecedent would lead from true premises to a false conclusion. As they also note, however, while abduction and induction are not *guaranteed* to preserve truth, these forms of inference may still be taken to be highly reliable guides to truth. At least we trust them in daily practice to preserve truth with high probability. Insofar as this practice is justified, modus ponens is still a highly reliable inferential principle. Put differently, it may still be near-valid in the sense discussed earlier in relation to McGee's version of Stalnaker's semantics.

To be sure, much about inferentialism is still unknown. What does it imply about the probabilities of conditionals? Which inferential principles beyond the ones just mentioned does it validate? Also, the proposal by Krzyżanowska and coauthors may need refinement. It was previously said that, for a variety of reasons, it may be necessary to go beyond the classical truth values ("true" and "false") and allow for the possibility that a sentence is neither true nor false. There is ample evidence that people sometimes evaluate conditionals as lacking a truth value (see also below). If inferentialism is to be descriptively adequate, it will have to accommodate this evidence, and so will have to be extended to incorporate indeterminacy as well. One seemingly natural proposal is to stipulate that a

---

relevance conditionals (Bennett, 2003). Krzyżanowska and coauthors explicitly propose their brand of inferentialism as a semantics for standard conditionals, not for unconditionals.

conditional is true if there is a strong enough argument from antecedent (plus background knowledge) to consequent; false if either there is an argument connecting antecedent and consequent but the argument is very weak, or there is an argument (perhaps only a weak one) from the antecedent to the negation of the consequent; and neither true nor false if there is no argument from antecedent to consequent at all. That, at any rate, is what we proposed in Douven et al. (2018).

Be this as it may, in the following we shall be concerned with inferentialism as it stands. Already as it stands, inferentialism makes some clear predictions about how people will evaluate particular conditionals. More importantly, it occurred to us that inferentialism makes clear predictions about the sort of data that we had gathered in the context of the work reported in Douven et al. (2018). As noted previously, inferentialism is an essential component of HIT, but our focus in Douven et al. (2018) was on the psychological theory rather than on the descriptive adequacy of the semantic theory. Thus, we did not earlier try to establish whether the data from our previous work are in accordance with inferentialism. That is what we aim to do here. Before turning to the new analysis, we first review Experiment 1 from Douven et al. (2018), in which the data were collected.

### 3 Experiment

In Douven et al. (2018), we introduced HIT, which abbreviates “Hypothetical Inferential Theory.” Taking its cue from the conjunction of inferentialism and dual-process theory, HIT offers a novel psychological account of conditionals, that is, of how people make sense of conditionals. Its main tenets, motivated by the aforesaid conjunction, are (i) that the default relevant mental representation of the conditional is the inferential relation between antecedent and consequent, and (ii) that this inferential link is strong, but perhaps not maximally strong—it just has to be *strong enough* (as judged from the interpreter’s perspective).

From these tenets, we derived a number of general predictions, the most notable ones being that parameters affecting the strength of the inference from antecedent to consequent should predict a conditional’s likelihood to be judged true, and that evaluations of conditionals should display response patterns analogous to those identified for inferences. In particular, we predicted and found a belief bias analogue. Belief bias (Evans, Barston, & Pollard, 1983) is the way that inferences are affected by the believability of their contents (usually the content of the conclusion). If the consequent of a conditional is analogous to an inference’s conclusion, we should find that the believability of the consequent affects belief in the conditional as a whole. This is exactly what we found in Douven et al. (2018).

As Nickerson, Barch, and Butler (2019) note, at least part of the “gappiness” we sense in conditionals like (1a) and (1b) may be due to the fact that their component parts concern seemingly unrelated topics. A conditional such as

- (5) If Mozart died from food poisoning, he began composing at the age of 3.

strikes us as gappy as well, but probably to a somewhat lesser extent than the aforementioned conditionals. While one is hard pressed to conceive of a plausible inferential connection between the antecedent and the consequent of (5), at least they pertain to the same person. Because in our previous research the main focus was on the role inferential connections played in the representation of conditionals, and not on topic coherence (which may merit a separate study), we decided to use as stimuli conditionals whose antecedents were at least semantically clearly related to their consequents, in that the antecedents and consequents of all conditionals that were used had the same well-delineated subject matter. Specifically, we used conditionals pertaining to a series of color patches in



Figure 1: Soritical color series that was the subject matter of the conditionals used in Experiment 1 from Douven et al. (2018).

which the color from one patch to the next varies gradually from clearly blue to clearly green. The conditionals the participants were asked to judge were all to the effect that if a particular patch in the series is green (blue), then a certain different patch in the series is green (blue) as well.<sup>8</sup>

### 3.1 Method

*Participants.* There were 704 participants in Experiment 1 from Douven et al. (2018), with 532 remaining after selection on the basis of native language, the passing of attention checks, and a color vision test. All participants were financially compensated for their time and effort spent on the survey. The mean age of the participants retained for the analysis was 34 ( $\pm 13$ ).

*Design, materials, and procedure.* We used a 3 (presentation condition: series stayed in sight/series did not stay in sight/series not shown but description given)  $\times$  2 (color condition: blue/green)  $\times$  2 (spread condition: large/small) between-participants design, resulting in 12 groups to which participants were assigned at random upon entering the survey.<sup>9</sup>

The presentation condition split the participant pool into three groups. Participants in two of these groups were shown the “soritical” color series seen in Figure 1; the remaining participants were not shown the series but were presented with the following text:

Imagine a series of 14 color patches, numbered 1 through 14, and ordered from left to right. The series begins with a clearly blue patch—patch number 1—on the left. The patches then gradually become more greenish as we progress to the right, with adjacent patches being almost indistinguishable in color. The series ends with a clearly green patch—patch number 14—on the right.

Of the two groups that were shown the series, one (the out-of-sight group) saw it only at the beginning of the experiment, while the other (the in-sight group) could look at the series throughout the experiment.

All participants were asked to evaluate 22 conditionals, each being of the schematic form

(6) If patch number  $i$  is  $X$ , then so is patch number  $j$ .

with  $i \in \{2, 7, 8, 9, 10, 13\}$  for all participants, but  $X$  either “blue” or “green,” depending on whether the participant was in the blue or in the green condition (this split was made purely for control reasons), and  $j$  depending on the so-called spread condition the participant was in, which determined which consequents were presented together with the six antecedents: in the small condition, the distance between the antecedent patch and the consequent patch was either 1 or 2 steps in each direction, while in the large condition, it was either 1 or 3 steps. The fact that steps of minus 2/3 from 2 and plus 2/3 from 13 were not possible for our soritical series from 1 to 14 explained why participants were asked to evaluate 22 conditionals (i.e.,  $4 \times 4 + 2 \times 3 = 22$ ).

<sup>8</sup>Note that, as a result, the conditionals’ component parts also had the same term specificity, in the sense of Gazzo Castañeda and Knauff (2019), which these authors showed to matter to how conditionals are evaluated.

<sup>9</sup>For data on group sizes, see Table 6.1 in Douven et al. (2018, p. 57).

The participants were given three response options to judge the truth of each conditional: “True,” “False,” and “Neither true nor false.” For the participants in the two “visual” groups (the out-of-sight and in-sight groups), the experiment had a second part, in which they were asked to indicate of each of the fourteen patches whether it was blue, green, or borderline blue/green.

To see the relevance to HIT of the materials that were used, it is to be noted that with each of the conditionals we can associate an argument, where the strength of the argument depends on the patches referred to in the conditionals. For instance, if we move to the right in the color series shown in Figure 1, the patches become greener. Thus, with

(7) If patch number 8 is green, so is patch number 11.

we can associate the argument that because patch number 11 is to the right of patch number 8, the supposition that patch number 8 is green allows us to infer that—a fortiori, as one might add—patch number 11 is green as well. Similarly, given the soritical nature of the color series, we can associate with

(8) If patch number 8 is green, so is patch number 7.

the argument that because adjacent patches are very similar in color, from the assumption that patch number 8 is green we can infer that patch number 7 is green as well. Importantly, the argument associated with (7) appears stronger than the one associated with (8). That is because in the case of the latter there is a countervailing consideration, to wit, that patches get *less* green when we move to the left, while in the former case there is no such consideration.

This highlights the importance of “direction”—to which side of the antecedent patch the consequent patch is—to argument strength. What in Douven et al. (2018) was called “distance”—how far removed from the antecedent patch the consequent patch is—is a further factor. For example, that immediately adjacent patches are similar in color is a more plausible contention than that patches two or three steps removed from each other are similar in color. As a result, we are inclined to deem the argument associated with (8) more compelling than the otherwise similar argument that we can associate with

(9) If patch number 8 is green, so is patch number 5.

In short, in our materials we manipulated strength of inferential connection by varying both direction and distance.

### 3.2 Results

As mentioned previously, we derived from the main tenets of HIT the prediction that strength of inferential connection between a conditional’s component parts determined the likelihood that the conditional is judged as true. For our materials, that translated to the more specific prediction that this likelihood is determined by direction and distance. An analysis of the data using a binomial generalized linear-mixed model revealed both to be highly reliable predictors of truth judgments indeed, thereby supporting HIT. For instance, we found that, as expected, (7) was reliably deemed true by participants, and more reliably so than (8), which in turn was more reliably deemed true than (9).

Another prediction we made was that the evaluation of conditionals would exhibit the same biases that are known to affect the evaluation of arguments, most notably a belief bias. We found strong evidence indeed for the presence of a belief bias in the data. In particular, we found consequent patch “rank” (i.e., where in the series the patch referred to in the consequent is to be found) to be

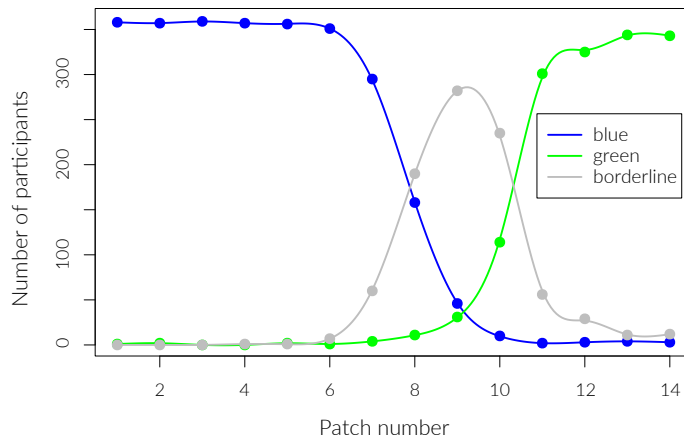


Figure 2: For each patch, the number of participants that classified the patch as blue (blue line), green (green line), and borderline blue/green (gray line). (Intermediate values have been interpolated.)

a significant predictor of truth judgment. In the green condition, conditionals whose consequent patch was toward the right (green) end of the color series were significantly more likely to be judged true than those whose consequent patch was further removed from that end, all else being equal. The opposite held for the blue condition. In particular, there was an unmistakable asymmetry in how the shades in the series shown in Figure 1 were judged by the participants in the visual groups. Figure 2 plots the response counts for each patch, showing that while patches 1–6 were judged to be blue by virtually all participants who saw the patches, only patches 13 and 14 were judged to be green by almost the same numbers of those participants. This asymmetry was reflected in the different belief bias patterns exhibited by the two visual groups. Furthermore, participants in the description group were also found to be susceptible to belief bias, but the bias in their responses did not show any asymmetry—which would have been a puzzling find, given that there was no asymmetry in how the series was described to them.

#### 4 Reassessing the semantics

As mentioned, the participants in the visual groups were, in a second part of Experiment 1 from Douven et al. (2018), asked to judge the color of each patch separately, where their response options were “blue,” “green,” and “borderline blue/green.” These responses, graphically summarized in Figure 2, were used in the aforementioned paper to investigate the presence of a belief bias in the truth judgments of the conditionals, as explained above. In this section, we want to leverage the same responses to shed new light on the semantics discussed in Section 2, although we will also, toward the end of the section, draw on the full set of data from Experiment 1 in Douven et al. (2018).

To bring the data from the visual groups in that experiment to bear on the semantics of conditionals, we assume that if a participant judged patch number  $i$  to be  $X$ , with  $X \in \{\text{blue, green}\}$ , then he or she would also judge “Patch number  $i$  is  $X$ ” to be true. Moreover, virtually all philosophers who have been concerned with vagueness hold that borderline cases give rise to indeterminacy in truth value (see, e.g., Schiffer, 2003, Ch. 5, and references given there). Accordingly, we assume that a participant would judge “Patch number  $i$  is  $X$ ” to be neither true nor false if he or she judged patch number  $i$  to be borderline blue/green.

Given these assumptions, we can analyze the frequencies with which the semantics at issue were violated by the participants. We first analyze the frequencies of violations of the main extant semantics—so all semantics discussed in Section 2, minus inferentialism—using the data from the visual groups. That only these data are used here is because we need the responses to the color classification task to classify the conditionals’ antecedents and consequents as true, false, or neither true nor false.

As noted in Douven et al. (2018), the data from Experiment 1 in that paper are not at all supportive of non-propositionalism, according to which conditionals are never true or false: “neither true nor false” responses constituted only 10.5 percent of the total responses to the conditionals questions, where it was further noted that 55 percent of all participants *never* chose the “neither true nor false” option. In fairness, non-propositionalists could claim that, insofar as participants chose the “true” or “false” option, they were using these terms pleonastically. While in that case the data would still not *support* non-propositionalism—for that, the position would have to come with a set of rules for when to use the truth and falsity predicates pleonastically, which is not the case—neither would the data undermine non-propositionalism.

Frequency counts for the other semantics require more preparation. First consider the three-value view, according to which a conditional has the semantic value of its consequent if its antecedent is true and else is neither true nor false. Recall that the standard version of the three-value view only pertains to conditionals whose antecedent and consequent have a determinate truth value. So, still assuming that if a participant in the green condition judged patch number  $i$  to be green/blue/borderline, he or she would judge “Patch number  $i$  is green” to be true/false/neither true nor false, respectively, and similarly for participants in the blue condition, participants can violate the three-value view in three possible ways, given our materials. For a conditional of the form “If patch number  $i$  is  $X$ , so is patch number  $j$ ,” with  $X$  either green or blue, a participant can violate it by judging patch number  $i$  to be  $Y$  (with  $Y$  blue if  $X$  is green, and vice versa) while not judging the conditional to be neither true nor false; by judging both patch number  $i$  and patch number  $j$  to be  $X$  yet not judging the conditional to be true; and by judging patch number  $i$  to be  $X$ , patch number  $j$  to be  $Y$ , while not judging the conditional to be false. Letting  $PQR$  indicate that the antecedent of a given conditional is evaluated as  $P$ , the consequent as  $Q$ , and the conditional as  $R$ , with  $P$ ,  $Q$ , and  $R$  being variables ranging over the three semantic values, the first type of violation consists of all combinations of judgments FTT, FFT, FTF, and FFF, the second of all TTF and TT# combinations of judgments, and the third of all TFT and TF# combinations of judgments.

Like the three-value view, the material conditional account is silent on what to say about a conditional in case the antecedent or consequent is evaluated as neither true nor false. So then, given our materials, there are two ways to violate this account, to wit, first, by not judging “If patch number  $i$  is  $X$ , then so is patch number  $j$ ” false while judging patch number  $i$  to be  $X$  and patch number  $j$  to be  $Y$ —which amounts to the TFT and TF# combinations of judgments—and second, by not judging that conditional true while either judging patch number  $i$  to be  $Y$  or judging patch number  $j$  to be  $X$  (or both), which are the TTF, TT#, FTF, FT#, FFF, and FF# combinations of judgments.

Stalnaker’s semantics might seem to pose more of a problem for the kind of analysis we are undertaking here, given that it is not truth-functional. To repeat, according to this semantics a conditional has the truth value its consequent has in the actual world, if its antecedent is true, and it has the truth value its consequent has in the nearest possible world in which its antecedent is true, if its antecedent is false in the actual world. Exactly what this implies concerning the materials from Douven et al. (2018) depends on how the nearness relation between worlds gets spelled out. Some, including Stalnaker, prefer to leave this relation fairly unconstrained, imposing only some



obvious logical strictures on it, such as, most notably, that if  $\phi$  is true in a world, that world is its own nearest  $\phi$ -world. On this construal, Stalnaker's semantics is predictively much weaker than the other semantics considered here. That is not a critique if the semantics is thought of as a tool for modeling particular linguistic phenomena—as it is in the eyes of many philosophers and linguists. But even in this case, it is predictively not altogether empty. Most importantly for our present concerns, it still predicts that people will judge a conditional to be true whenever they judge both its antecedent and its consequent to be true, that, in other words, their judgments will respect Centering; that is a direct consequence of the aforementioned fact that a world in which  $\phi$  holds true is its own nearest  $\phi$ -world.

Others have followed Lewis (1979) in assuming a less liberal explication of nearness between worlds. On Lewis' proposal, nearness is a matter of agreement in laws of nature and agreement in particular fact, where agreement in laws of nature is to weigh more heavily, and particular facts are to be understood as basic, non-supervenient facts. Often it is no easy matter to determine nearness even given these criteria. We lack formal means for weighing against each other different disagreements in laws of nature and/or in particular fact. However, because the possible worlds we need to consider for our purposes are of an exceedingly simple structure, Lewis' proposal does make it easy to determine nearness relations between them.

Suppose the patch referred to in the consequent of a conditional, but not that referred to in the antecedent, has color  $X$ , at least in the perception of a participant. Then a possible world in which the antecedent and consequent patches are both  $X$  is closer to (what the participant takes to be) the actual world than a world in which the antecedent patch is  $X$  but the consequent patch is not: the former world is in agreement with the actual world on one particular fact more than the latter—which settles the matter, given that laws of nature do not even enter the picture in the possible worlds at issue in the present context. Thus, for instance, the conditional (9) should—according to Stalnaker's semantics—be evaluated as true by any participant who deems the consequent patch blue, whatever color the participant deems the antecedent patch to have. More generally, any conditional whose consequent is evaluated as true should be evaluated as true.

So far, Stalnaker's semantics (assuming Lewis' explication of nearness) agrees with the material conditional account, when confined to our materials. The agreement does not extend to the case in which, in the participant's perception, neither the antecedent patch nor the consequent patch is of the color that the antecedent and the consequent, respectively, attribute to it. Consider again a concrete example:

(10) If patch number 2 is green, so is patch number 3.

In the actual world, patch number 2 and patch number 3 are both clearly blue. Which is closer to the actual world: a possible world in which both patches are green, or one in which patch number 2 is green and patch number 3 is blue? The latter, obviously, given that, in it, patch number 3 has the same color it has in the actual world.<sup>10</sup> Hence, (10) is false: its antecedent is false, and in the closest world in which its antecedent is true, its consequent is (still) false. More generally, a conditional whose consequent is evaluated as false should be evaluated as false. This is so even if its antecedent is false, in which case the conditional would be true according to the material conditional account.

---

<sup>10</sup>It might be thought that the former world preserves the fact that the two patches are of the same color. But that fact supervenes on the facts concerning the colors of the patches, and is therefore not to be considered in determining closeness between worlds. This relies on Lewis' aforementioned proposal, and—as an anonymous referee pointed out—there is no clear consensus on how to define closeness, and different proposals lead to different predictions. However, Lewis' proposal does have many advocates in philosophy and we ourselves deem it plausible enough to assume it here.

It thus appears that, even though the reference to non-actual possible worlds makes Stalnaker’s semantics non-truth-functional in general, the semantics *is* truth-functional given the materials from Douven et al. (2018), at least assuming Lewis’ understanding of nearness between worlds: possible worlds in which the consequent patch has the color it has in the actual world are always closer to the actual world than possible worlds in which the consequent patch has a different color. As a result, the evaluation of the conditionals in the experiment ought to follow the evaluation of their consequent. Given that this semantics, too, says nothing about conditionals whose antecedent or consequent is evaluated as neither true nor false, we can discern two types of violations: judging “If patch number  $i$  is  $X$ , then so is patch number  $j$ ” false or neither true nor false while judging patch number  $j$  to be  $X$ , which are the TTF, TT#, FTF, and FT# combinations of judgments; and judging the conditional true or neither true nor false while judging patch number  $j$  to be  $Y$ , which are the TFT, TF#, FFT, and FF# combinations of judgments.

We checked the frequencies with which all of the aforementioned patterns occur in the data. There was a totality of 3960 responses that register a participant’s evaluation of a conditional whose antecedent and consequent patches have been classified either as blue or as green by the participant. Of these responses, 2202, or 56 percent, violate the three-value view; 1604, or 41 percent, violate the material conditional account; and 1210, or 31 percent, violate Stalnaker’s semantics.

While the above semantics only consider conditionals with determinate constituents, we saw that some extensions of the three-value view pertain to conditionals with indeterminate constituents as well, in particular, the de Finetti table, the Farrell table, and the Cooper table. Each of these is violated by eighteen combinations of judgments, to wit, for each of the nine rows in Table 1, the combination of judgments that consists of the semantic value of the antecedent, the semantic value of the consequent, and one of the two semantic values different from the value of the conditional that the row gives in that table. Of the total of 7898 responses, 5750, or 73 percent, violate the de Finetti table; 5136, or 65 percent, violate the Farrell table; and 4609, or 58 percent, violate the Cooper table.

Table 2 summarizes the foregoing by listing all possible combinations of judgments, indicating whether or not they violate any of the main semantics, and giving the frequencies with which the combinations occurred in the data. Triples of combinations that concur in their antecedent and consequent judgments (henceforth called “sets of premises”) are separated by dashed lines.

Next, we evaluated the predictions presented in Table 2 across the main semantics by means of statistical modeling using MPTinR (Singmann & Kellen, 2013).<sup>11</sup> For this, it is important to note that, for each set of premises, each of the main semantics predicts exactly one response, if any (e.g., the three-value view predicts T for the TT case, the material conditional account predicts T for that case, and so on). Hence, one can compare the predictions with the observed frequencies per set of premises to assess the accuracy of the predictions. Given the variance across participants as documented in Douven et al. (2018), we planned to analyze the frequencies of violations on a per participant level. However, for only 28 of the 359 participants at least one response fell in each of the nine different sets of premises. When considering only the sets of premises with determinate premises, for only 90 participants at least one response fell in each of the four different sets of premises. Consequently, for participants with no responses for one or more of the sets of premises, we only tested those predictions for which they actually provided responses. Taken over all sets of premises (i.e., including sets with indeterminate premises), participants provided responses for between three and nine of those sets, with a median of providing responses for seven sets (mean = 6.8). When only considering the four sets of premises with determinate premises, participants provided responses for between one and four sets, with a median of providing responses for two sets of premises (mean = 2.7). For each

<sup>11</sup>The R scripts are available at: <https://osf.io/eskw5/>.

Table 2: Classification of various combinations of judgments according to the different semantics; red crosses indicate violations, green checkmarks indicate non-violations. The frequency column gives the frequencies with which the combinations occurred in the data.

	non-propositionalism	three-value view	material conditional	Stalnaker's semantics	de Finetti	Farrell	Cooper	frequency
T T T	X	✓	✓	✓	✓	✓	✓	1483
T T #	✓	X	X	X	X	X	X	91
T T F	X	X	X	X	X	X	X	266
T # I	X	-	-	-	X	X	X	144
T # #	✓	-	-	-	✓	✓	✓	58
T # F	X	-	-	-	X	X	X	300
T F T	X	X	X	X	X	X	X	28
T F #	✓	X	X	X	X	X	X	18
T F F	X	✓	✓	✓	✓	✓	✓	159
# T T	X	-	-	-	X	X	X	600
# T #	✓	-	-	-	✓	✓	X	73
# T F	X	-	-	-	X	X	X	234
# # I	X	-	-	-	X	X	X	513
# # #	✓	-	-	-	✓	✓	✓	131
# # F	X	-	-	-	X	X	X	534
# F T	X	-	-	-	X	X	X	93
# F #	✓	-	-	-	✓	✓	X	86
# F F	X	-	-	-	✓	✓	✓	700
F T T	X	X	✓	✓	X	X	X	103
F T #	✓	✓	X	X	✓	✓	✓	8
F T F	X	X	X	X	X	X	X	80
F # I	X	-	-	-	X	X	X	225
F # #	✓	-	-	-	✓	✓	✓	42
F # F	X	-	-	-	X	X	X	205
F F T	X	X	✓	X	X	X	X	611
F F #	✓	✓	X	X	✓	✓	✓	108
F F F	X	X	X	✓	X	X	X	1005

Table 3: Comparison of prediction accuracy across semantics.

non-propositionalism	three-value view	material conditional	Stalnaker's semantics	de Finetti	Farrell	Cooper
<i>all premises</i>						
1.0 (1.0)				.994 (.994)	.997 (.997)	.997 (.997)
42,187				32,211	28,369	25,023
<i>only determinate premises</i>						
.997 (1.0)	.92 (.95)	.72 (.84)	.61 (.72)	.92 (.95)	.92 (.95)	.92 (.95)
21,720	11,744	7,256	5,480	11,744	11,744	11,744

*Note.* The first row below each subheading shows the proportion of participants for which multinomial models representing probabilistic versions of the predictions of the semantics are rejected by the data with  $p < .005$  ( $p < .05$  in parentheses). The second row shows the summed  $G^2$  values across all participants. Truth and falsity of the premises was determined in the way explained in the text. Predicted responses were restricted to occur at least 95 percent of the time.

participant and subset (i.e., responses from one participant to one set of premises), the number of responses was ranging from one to sixteen, with a median of three responses per participant and set.

A common way for estimating the accuracy of multinomial predictions is via the  $G^2$  measure of goodness of fit (Agresti, 2002). Unfortunately,  $G^2$  (and similar measures like  $\chi^2$ ) cannot adequately deal with deterministic predictions (i.e., predictions of exactly 0 and 1) but only with stochastic predictions, that is, predictions within the (0, 1)-interval. To circumvent this problem, we employed a strategy used in Rieskamp and Otto (2006) (see also Rieskamp, 2008) to assess the adequacy of a deterministic decision heuristic, namely, allowing a certain amount of “application errors.” While Rieskamp and Otto estimated this amount from the data, we set it to their observed mean value of 5 percent per set of premises. For our data, this meant that, for each participant and each prediction following from each of the main semantics, we estimated for all the sets of premises for which the participant provided data, via maximum likelihood estimation, a saturated model (i.e., two parameters per set of premises) with the restriction that the predicted response was given at least 95 percent of the time (in other words, a different response than the predicted one in 5 percent of the cases would still provide a perfect fit to the data).

Although in line with previous research (Rieskamp & Otto, 2006), allowing for exactly 5 percent of “application errors” was a somewhat ad hoc choice. To alleviate concerns one might have about this, we decided to use a significance criterion of  $p = .005$  in the following. Note that in the present case a significance criterion with a lower threshold corresponds to a more lenient criterion (as we are interested in testing if participants conform to a specific semantics). That is, only if the observed misfit for a specific participant and semantics corresponds to a  $p$ -value of less than .005 do we count this as a violation of this semantics. Using a larger and therefore more stringent  $p$ -value threshold (such as .05) would have led us to observe even more violations as reported below.

Because each of the models corresponding to the different semantics is effectively imposing inequality restrictions on the multinomial distribution, misfits do not follow a  $\chi^2$  distribution but a mixture of  $\chi^2$  distributions (e.g., Self & Liang, 1988), also known as  $\tilde{\chi}^2$ . For example, the most conservative  $\tilde{\chi}^2$  distributions for individual misfits across four and nine sets of premises have critical values ( $p < .005$ ) of 11.50 and 17.04, respectively. We calculated the appropriate  $\tilde{\chi}^2$  distributions for each participant individually, depending on the number of sets of premises for which data was available.<sup>12</sup>

Results of this analysis showing the proportion of participants for which the probabilistic versions of the semantics are rejected ( $p < .005$ ) and summed  $G^2$  values are given in Table 3. Overall, the

<sup>12</sup>We are grateful to David Kellen for his advice regarding this analysis.

performance of the semantics in describing the response frequencies is extremely poor. Although, as expected from the previous analysis, non-propositionalism did worst in terms of summed  $G^2$ , the accounts by de Finetti, Farrell, and Cooper, as well as the standard three-value view were also rejected for more than 90 percent of the participants. The comparatively best account was provided by Stalnaker’s semantics with “only” 61 percent of the participants rejected. The material conditional account did second best with 72 percent of the participants rejected.

It is noteworthy that the second most frequent combination of judgments is FFF. (The FFF pattern is highly unusual in the psychological literature on conditionals; we come back to this in Section 5.) The 1005 FFF responses constitute 13 percent of the 7898 combinations of judgments we are considering, and 25 percent of the 3960 combinations of judgments that remain from leaving out the ones that attribute an indeterminate value to either the antecedent or the consequent (or both). That Stalnaker’s semantics, alone among the main semantics, countenances these combinations of judgments is probably one of the reasons why it performed comparatively well in the previous analysis and a clear point in favor of that semantics. From the perspective of inferentialism, the interesting fact about these 1005 FFF combinations of judgments is that 699 of them concerned what in Douven et al. (2018) were called “incongruent conditionals”—conditionals whose consequent patch lies in the less- $X$  direction, with  $X \in \{\text{blue, green}\}$ —and 306 concerned “congruent conditionals,” that is, conditionals whose consequent patch lies in the more- $X$  direction. A goodness of fit test on the summed frequencies showed this difference to be significant:  $\chi^2(1) = 153.681, p < .0001$ . This difference is precisely as one would expect, assuming inferentialism, given that arguments from antecedent to consequent may be expected to be perceived as stronger if the consequent patch is to the more- $X$  side of the antecedent patch than if it is to the less- $X$  side.

Having come to inferentialism now, we start by noting that, to make our data bear on this semantics, we must proceed differently than above. From the perspective of inferentialism, it is immaterial how a participant evaluates a conditional’s antecedent or consequent, whether in the actual world or in nearby worlds. What matters is whether, in light of the background knowledge, the antecedent can support a strong enough argument for the consequent. As mentioned in Section 2.5, what counts as strong enough is vague, but rather than a drawback of the semantics, this may be a virtue in that it may account for why it can be vague whether a conditional is true. Specifically with respect to our data, it may account for why minority responses on some conditionals were substantial: different individuals may draw the line for what counts as a strong enough argument differently. This makes it difficult to designate any single judgment on a conditional as a violation of inferentialism.

If inferentialism does not make definite predictions about each of the conditionals in our experiment, it does make clear predictions about which *patterns* are and are not to be found in the responses of participants. Most notably perhaps, and in contrast with the other semantics, inferentialism predicts not only violations of Centering but also a certain pattern in those violations.

As to the former, that a conditional has true component parts does not entail that there is a strong enough argument connecting these parts (whether or not with help from background premises), and surely not all conditionals in our materials can be associated with a strong argument connecting their antecedent and consequent, given any reasonable reading of “strong enough.” Hence, from the perspective of inferentialism it is unsurprising that, as can be quickly seen on the basis of Table 2, we find a violation of Centering in one fifth of the cases in which a conditional is judged to have a true antecedent and consequent.

As to what pattern to expect in those violations, inferentialism predicts that such violations are more frequent for incongruent conditionals than for congruent conditionals. And we find indeed that of the TT responses (responses in which antecedent and consequent are judged true) for incongruent conditionals, 32 percent violate Centering. By contrast, of the TT responses for

congruent conditionals, only 10 percent violate Centering. This difference is highly significant for the summed frequencies:  $\chi^2(3) = 917.6, p < .0001$ .<sup>13</sup>

A further pattern that inferentialism predicts to be present in the data needs more explaining. Note that inferentialism does not just suggest that direction, distance, and rank of consequent patch play a role in predicting how participants evaluate the various conditionals. It also implies a kind of consistency requirement, to wit, that conditionals with relatively stronger corresponding arguments be evaluated as true if conditionals with relatively weaker corresponding arguments are evaluated as true, and that conditionals with equally strong (or equally weak) corresponding arguments be evaluated alike.

Even without a formal account of generalized validity (in particular, without generally accepted formal accounts of abductive and inductive validity) and a generally accepted precise measure of argument strength, we have no difficulty comparing at least some of the arguments we naturally associate with our stimuli with respect to their strength. For example, for instances of the schema

(ii) If patch number  $i$  is  $X$ , then so is patch number  $i + j$ .

with  $X$  green, it seems uncontentious to say that, for fixed  $i$ , the corresponding argument is stronger if  $j = -1$  than if  $j = -2$  or  $j = -3$  as the last two are further away on the incongruent side from  $i$  than  $j = -1$ . *Mutatis mutandis*, the same holds for the blue condition. Also, the arguments corresponding to the instances with  $j = 1, j = 2$ , or  $j = 3$  seem equally strong, and they are clearly stronger than the arguments for the instances with  $j = -1, j = -2$ , or  $j = -3$ , as in the green condition positive  $j$ s are on the congruent side and negative  $j$ s on the incongruent side.

So, someone who judges the instances of (ii) with  $X$  green and  $j \in \{1, 2, 3\}$  to be true is not inconsistent if he or she judges the instances with  $X$  green and  $j \in \{-1, -2, -3\}$  to be false or neither true nor false; he or she may think that the arguments corresponding to the latter instances are not strong enough. If, on the other hand, a participant judges the instance with, say,  $j = -2$  to be true, but not all, or even not any, of the instances with  $j > -2$ , then that participant *is* inconsistent.

It is difficult to go beyond these observations and say in general how the conditionals are to be ordered in terms of strength of corresponding arguments. There is, for example, no formal basis for regarding the argument corresponding to

(12) If patch number 2 is green, so is patch number 1.

as being weaker or stronger than the argument corresponding to

(13) If patch number 7 is green, so is patch number 4.

Generally speaking, the problem is that the considerations that are at play in this kind of argument may pull in different directions, while the background knowledge available to the participants does not provide any information on which consideration is to take precedence in the case of conflict (if any is to take precedence at all).

But the observation made two paragraphs back provides ample opportunity for the further testing of inferentialism. To this end, we again considered the full data set of all 532 participants to Experiment 1 from Douven et al. (2018) and, for each participant individually, separated the trials into subsets for each antecedent  $i$ . In each of those subsets we searched for the trial with the “most incongruent  $j$ ” (i.e., the lowest  $j$  in the green condition and the largest  $j$  in the blue condition) for which the response was “True.” Based on this trial, we only considered trials with “more congruent  $j$ s” (if any existed) and counted all further “True” responses as consistent and all other responses as violations. Table 4 gives an overview of the number of congruent responses and violations summed

<sup>13</sup>Needless to say, given Inferentialism, one expects to find many responses *in accordance with* Centering. For example, it is reasonable to suppose that anyone with normal vision will judge patches 13 and 14 to be green. And Inferentialism also makes it reasonable to expect anyone to judge as true the conditional “If patch number 13 is green, so is patch number 14.”

Table 4: Frequency and proportion of violations for inferentialism.

antecedent	violations	congruent	proportion
2	54	522	.09
7	140	550	.20
8	139	515	.21
9	133	496	.21
10	108	559	.16
13	73	464	.14
sum	647	3106	.17

*Note.* Construction of this table is described in the main text.

across the 524 participants for which we found corresponding trials. What is obvious from this table is that, although there is a considerable number of violations, the congruent responses are clearly the majority.

To further test these data, we performed an analysis analogous to the one reported above for the main semantics and fitted the individual frequencies per antecedent with saturated multinomial models representing the predictions of inferentialism in a probabilized form (i.e., predicting at least 95 percent congruent responses per participant  $\times$  antecedent combination). Note that, similar to the frequency data of the main semantics, participants did not always provide data for all six antecedents; the range was providing data from one to six antecedents with a median of providing data for five antecedents (mean = 4.7). Hence, we again calculated the appropriate  $\chi^2$  distributions for each participant individually, depending on the number of antecedents for which participants provided data. This analysis revealed that only for 19.1 percent of the participants were the predictions of inferentialism rejected by the data at  $p < .005$  (27.1 percent at  $p < .05$ ). The overall summed  $G^2$  was 2,871.

Although these results are not *formally* comparable to the data for the main semantics, as the data used to fit the models is different (and so we cannot directly compare the  $G^2$ -values across semantics), they nevertheless provide a strong argument in favor of inferentialism when considering the absolute size of violations. More specifically, the degree to which a specific semantics is violated can be seen as the degree to which it fails a test for its support, which can be understood as a measure of effect size. For inferentialism, the proportion of participants which violate this semantics is small (less than 22 percent) relative to the proportions of violations we found for the main semantics (larger than 60 percent).

A justified question is whether this large difference in the proportion of participants that show violations could be due to factors unrelated to the predictions of the semantics. For example, recent research on uncertain inferences has controlled for the possibility that a specific response can be in line with theoretical predictions purely by chance (Singmann, Klauer, & Over, 2014; Evans, Thompson, & Over, 2015).<sup>14</sup> To perform a similar correction in the present case, we first need to calculate the probability that a specific response is in line with a semantics by chance. Following Evans et al. and Singmann et al., we assume that chance responses correspond to a uniform distribution on the response scale. Because for each semantics always one of the three response options was correct, the probability that any one observation is in line with a specific semantics by chance is therefore 1/3. However, in contrast to the work on uncertain inferences, we do not evaluate agreement with a specific semantics independently for each response, but jointly across all responses per participant.

<sup>14</sup>We thank an anonymous reviewer for suggesting to investigate this question.

Therefore, the by-participant probability for being in line with a semantics by chance is  $(1/3)^k$ , where  $k$  is the number of responses per participant that are relevant for testing a specific semantics. Because the chance probability decreases quickly with increasing  $k$ , and  $k$  is usually larger for the main semantics than for inferentialism, a part of the large difference in the proportion of participants that show violations could potentially be explained for by a higher chance of being in line with inferentialism than with the main semantics. As shown in the following, this difference in being correct by chance is by far not enough to explain the large difference in the proportion of participants which violate the different semantics.

When considering the main semantics and both determinate and indeterminate premises, each participant provided 22 responses corresponding to a chance probability of being in line with any such semantic of less than 0.001 percent. When only considering determinate premises, participants provided between 4 and 22 responses (mean = 11.0, median = 10). This means that the largest possible probability for being in line with any of the main semantics by chance is 1.2 percent (i.e.,  $(1/3)^4$ ). When considering the distribution of chance probabilities across participants for those semantics, the mean is 0.04 percent and the median is 0.002 percent. For inferentialism, participants provided between 1 and 16 responses (mean = 7.2, median = 6). Thus, the largest possible probability for being in line with inferentialism by chance was 33.3 percent and the smallest such probability was again less than 0.001 percent. When considering the distribution of chance probabilities across participants for inferentialism, the mean is 1.5 percent, and the median is 0.1 percent. Thus, on average it is highly unlikely that the large advantage of inferentialism is only due to more participants being in line with it by chance. More specifically, for only 17.0 percent of participants the probability to be in line with inferentialism was above 1 percent (for 2.1 percent the chance probability was 33.3 percent, for 3.1 percent it was 11.1 percent, for 6.1 percent it was 3.7 percent, and for 5.1 percent it was 1.2 percent). Even if those 17.0 percent of participants responded in exactly the manner consistent with inferentialism and we would also count their responses as violations of inferentialism, the overall proportion of participants that show violations would still be only 34 percent, roughly half of the smallest proportion of violations for the main semantics.

## 5 General discussion

In previous work, we proposed a psychological account of the interpretation of conditionals. Hypothetical Inferential Theory combined a dual-processing approach on the processing level with inferentialism on the computational level. This combination led us to hypothesize that the default mental representation of a conditional is the inferential connection between its component parts, where this connection needs only be “strong enough.” In that earlier work, we made no attempt to empirically investigate the descriptive adequacy of inferentialism as a semantics. Indeed, we intentionally remained noncommittal vis-à-vis inferentialism as a *semantics* of conditionals, leaving open the possibility that the requisite inferential connection might have a pragmatic origin, resulting from certain conventions governing language use.

In the present paper, we re-analyzed the data from Experiment 1 reported in Douven et al. (2018), now with an eye toward testing inferentialism as a semantics, that is, an account of the truth conditions of conditionals. We also specifically wanted to compare inferentialism with the more traditional semantics of conditionals. As we saw in the above, the various semantics make clear, and partly very different, predictions about the materials used in Douven et al. (2018), in particular about which patterns we *should*, and which we should *not*, find in the responses to those materials. In our new analysis, we verified to what extent these predictions were corroborated by the data.

In this analysis, non-propositionalism performed the worst, at least when responses were taken at face value (so as being genuine attributions of truth and falsity, insofar as the “truth” and “falsity” options were chosen). However, the other traditional semantics performed poorly (the material conditional account and Stalnaker’s semantics) to very poorly (the remaining ones) as well, with—for



most of the semantics—over 90 percent of the participants significantly violating the predictions. A further noteworthy finding in our data was a clear pattern of violations of Centering, which was exactly as one would expect given inferentialism but not given any of the standard semantics of conditionals, all of which *entail* the principle. Moreover, we saw that inferentialism predicts that, however the participants exactly evaluate the conditionals in our materials, they will satisfy certain consistency requirements, and that this prediction was violated by less than 20 percent of the participants. This difference in the proportion of participants violating predictions of the main semantics and predictions of inferentialism also cannot be explained by chance considerations.

While we take these results to support inferentialism, we cannot rule out that theorists will be able to develop a pragmatic explanation of our findings on an ad hoc basis. Here, we can only challenge those theorists to provide us with a worked-out such explanation, which at present is missing from the literature. *Mutatis mutandis* for the suggestion (which some might want to make) that semantics based on the Equation or the Ramsey Test may be able to take into account the notions of congruency, distance, and direction, which were central to the inferentialist explanation of the data we looked at: we would like to see the details of such a proposal.

Although the basic idea of inferentialism dates back to antiquity, in its present form the position is still in its infancy. As noted in Section 2.5, more needs to be said about what inferentialism implies about the probabilities of conditionals. One way to reconcile inferentialism with the known body of data supporting the Equation is to see the basic idea underlying inferentialism as filling in the details of the procedure in the footnote by Ramsey (1929/1990) that has been interpreted as suggesting the Ramsey Test mentioned in Section 2.2. That would also make sense of the fact that we find both that footnote and Ramsey's earlier-cited endorsement of Mill's version of inferentialism in the same paper. (Ramsey himself makes no effort to connect the footnote to the passage about Mill that was cited on page 3.) In the procedure described in the footnote, we add one proposition hypothetically to our stock of beliefs, make minimal adjustments to maintain consistency (if necessary), and then determine how probable a second proposition is in that hypothetical belief state. In reading the footnote, one cannot fail to notice that Ramsey is entirely silent on *how* we are to determine the probability of the second proposition in the belief state that results from hypothetically updating on the first. One suggestion (Douven, 2016a, chap. 2) is that this process is guided by an assessment of the strength of the argument from the first proposition to the second, given the rest of the hypothetical belief state: the stronger we take the argument to be, the higher we set the conditional probability (and the probability of the conditional). Naturally, whether assessing argument strength is really a good guide to assessing probabilities, and whether people do use it as such a guide, are questions that are open to empirical investigation. We only flag this interpretive option of the two passages in Ramsey (1929/1990) as an avenue for future research.

In conclusion, recent work in philosophy has shown that inferentialism has long been dismissed for the wrong a priori reasons. Our experiment shows that the position is not only a priori tenable but also enjoys empirical support. This makes it well worth the effort of looking further into the position.

## Acknowledgments

We are greatly indebted to David Over for many useful discussions on the subject matter of this paper. Thanks also to two anonymous referees and Christopher von Bülow for valuable comments on previous versions.

## References

Adams, E. W. (1965). The logic of conditionals. *Inquiry*, 8, 166–197.

- Adams, E. W. (1966). Probability and the logic of conditionals. In J. Hintikka & P. Suppes (Eds.), *Aspects of inductive logic* (pp. 265–316). Amsterdam: North-Holland.
- Adams, E. W. (1975). *The logic of conditionals*. Dordrecht: Reidel.
- Adams, E. W. (1998). *A primer of probability logic*. Stanford CA: CSLI Publications.
- Adams, E. W. & Levine, H. P. (1975). On the uncertainties transmitted from premises to conclusions in deductive inferences. *Synthese*, 30, 429–460.
- Agresti, A. (2002). *Categorical data analysis*. New York: Wiley-Interscience.
- Baratgin, J. & Politzer, G. (2016). Logic, probability and inference: a methodology for a new paradigm. In L. Macchi, M. Bagassi, & R. Viale (Eds.), *Cognitive unconscious and human rationality* (pp. 119–142). Cambridge MA: MIT Press.
- Baratgin, J., Politzer, G., Over, D. E., & Takahashi, T. (2018). The psychology of uncertainty and three-valued truth tables. *Frontiers in Psychology*, 9, 1479. <https://doi.org/10.3389/fpsyg.2018.01479>
- Bennett, J. (2003). *A philosophical guide to conditionals*. Oxford: Oxford University Press.
- Blamey, S. (1986). Partial logic. In D. M. Gabbay & F. Guenther (Eds.), *Handbook of philosophical logic, Vol. III* (pp. 1–70). Dordrecht: Kluwer.
- Braine, M. D. S. (1978). On the relation between the natural logic of reasoning and standard logic. *Psychological Review*, 85, 1–21.
- Braine, M. D. S. & O'Brien, D. P. (1991). A theory of *if*: Lexical entry, reasoning program, and pragmatic principles. *Psychological Review*, 98, 182–203.
- Burgess, J. P. (2004). Review of J. Bennett, *A philosophical guide to conditionals*. *Bulletin of Symbolic Logic*, 10, 565–570.
- Cruz, N., Baratgin, J., Oaksford, M., & Over, D. E. (2015). Bayesian reasoning with ifs and ands and ors. *Frontiers in Psychology*, 6, 192. doi: 10.3389/fpsyg.2015.00192
- Cruz, N., Over, D. E., Oaksford, M., & Baratgin, J. (2016). Centering and the meaning of conditionals. In A. Papafragou, D. Grodner, D. Mirman, & J. C. Trueswell (Eds.), *Proceedings of the 38th annual conference of the cognitive science society* (pp. 1104–1109). Austin TX: Cognitive Science Society.
- de Finetti, B. (1995). The logic of probability. *Philosophical Studies*, 77, 181–190.
- Douven, I. (2008). The evidential support theory of conditionals. *Synthese*, 164, 19–44.
- Douven, I. (2011). Indicative conditionals. In L. Horsten & R. Pettigrew (Eds.), *A companion to philosophical logic* (pp. 383–405). London: Continuum Press.
- Douven, I. (2016a). *The epistemology of indicative conditionals*. Cambridge: Cambridge University Press.
- Douven, I. (2016b). On de Finetti on iterated conditionals. In C. Beierle, G. Brewka, & M. Thimm (Eds.), *Computational models of rationality* (pp. 265–279). London: College Publications.
- Douven, I. (2017). How to account for the oddness of missing-link conditionals. *Synthese*, 194, 1541–1554.
- Douven, I., Elqayam, S., Singmann, H., & van Wijnbergen-Huitink, J. (2018). Conditionals and inferential connections: A hypothetical inferential theory. *Cognitive Psychology*, 101, 50–81.
- Douven, I. & Verbrugge, S. (2010). The Adams family. *Cognition*, 117, 302–318.
- Douven, I. & Verbrugge, S. (2013). The probabilities of conditionals revisited. *Cognitive Science*, 37, 711–730.
- Edgington, D. (1992). Validity, uncertainty and vagueness. *Analysis*, 52, 193–204.
- Edgington, D. (1995). On conditionals. *Mind*, 104, 235–329.
- Edgington, D. (1997). Vagueness by degrees. In R. Keefe & P. Smith (Eds.), *Vagueness: A reader* (pp. 294–316). Cambridge MA: MIT Press.
- Edgington, D. (2003). What if? Questions about conditionals. *Mind & Language*, 18, 380–401.
- Elqayam, S. & Over, D. E. (2013). New paradigm psychology of reasoning. *Thinking & Reasoning*, 19, 249–265.

- Evans, J. St. B. T. (2006). The heuristic–analytic theory of reasoning: Extension and evaluation. *Psychonomic Bulletin & Review*, 13, 378–395.
- Evans, J. St. B. T. (2007). *Hypothetical thinking: Dual processes in reasoning and judgement*. Hove: Psychology Press.
- Evans, J. St. B. T., Barston, J. L., & Pollard, P. (1983). On the conflict between logic and belief in syllogistic reasoning. *Memory & Cognition*, 11, 295–306.
- Evans, J. St. B. T., Handley, S. J., Neilens, H., & Over, D. E. (2007). Thinking about conditionals: A study of individual differences. *Memory & Cognition*, 35, 1759–1771.
- Evans, J. St. B. T., Handley, S. J., & Over, D. E. (2003). Conditionals and conditional probability. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 29, 321–355.
- Evans, J. St. B. T. & Over, D. E. (2004). *If*. Oxford: Oxford University Press.
- Evans, J. St. B. T. & Over, D. E. (2013). Reasoning to and from belief: Deduction and induction are still distinct. *Thinking & Reasoning*, 19, 267–283.
- Evans, J. St. B. T., Thomspon, V. A., & Over, D. E. (2015). Uncertain deduction and conditional reasoning. *Frontiers in Psychology*, 6, 398.
- Fugard, A., Pfeifer, N., Mayerhofer, B., & Kleiter, G. (2011). How people interpret conditionals: Shifts toward the conditional event. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 37, 635–648.
- Gazzo Castañeda, L. E. & Knauff, M. (2018). The specificity of terms affects conditional reasoning. *Thinking & Reasoning*, in press. doi: 10.1080/13546783.2018.1467343
- Gibbard, A. (1981). Two recent theories of conditionals. In W. Harper, R. Stalnaker, & G. Pearce (Eds.), *Ifs* (pp. 211–247). Dordrecht: Reidel.
- Gilio, A. & Over, D. E. (2012). The psychology of inferring conditionals from disjunctions: A probabilistic study. *Journal of Mathematical Psychology*, 56, 118–131.
- Gilio, A. & Sanfilippo, G. (2014). Conditional random quantities and compounds of conditionals. *Studia Logica*, 102, 709–729.
- Grice, H. P. (1989). Indicative conditionals. In his *Studies in the way of words* (pp. 58–85). Cambridge MA: Harvard University Press.
- Jackson, F. (1979). On assertion and indicative conditionals. *Philosophical Review*, 88, 565–589.
- Jackson, F. (1987). *Conditionals*. Oxford: Blackwell.
- Jeffrey, R. C. (1991). Matter of fact conditionals. *Proceedings of the Aristotelian Society* (Supplementary Volumes), 65, 161–183.
- Kneale, W. & Kneale, M. (1962). *The development of logic*. Oxford: Oxford University Press.
- Kratzer, A. (1986). Conditionals. In A. M. Farley & K. E. McCollough (Eds.), *Papers from the parasession on pragmatics and grammatical theory* (pp. 115–135). Chicago: Chicago Linguistics Society.
- Kratzer, A. (1991). Conditionals. In A. von Stechow & D. Wunderlich (Eds.), *Semantik: Ein internationales Handbuch zeitgenössischer Forschung/Semantics: An international handbook of contemporary research* (pp. 651–656). Berlin: de Gruyter.
- Kratzer, A. (2012). *Modals and conditionals*. Oxford: Oxford University Press.
- Krzyżanowska, K. H., Collins, P. J., & Hahn, U. (2017). Between a conditional’s antecedent and its consequent: Discourse coherence vs. probabilistic relevance. *Cognition*, 164, 199–205.
- Krzyżanowska, K. H., Wenmackers, S., & Douven, I. (2014). Rethinking Gibbard’s riverboat argument. *Studia Logica*, 102, 771–792.
- Lewis, D. K. (1975). Adverbs of quantification. In E. Keenan (Ed.), *Semantics of natural language* (pp. 169–218). Dordrecht: Reidel.
- Lewis, D. K. (1976). Probabilities of conditionals and conditional probabilities. *Philosophical Review*, 85, 297–315.
- Lewis, D. K. (1979). Counterfactual dependence and time’s arrow. *Noûs*, 13, 455–476.
- Lewis, D. K. (1981). Ordering semantics and premise semantics for counterfactuals. *Journal of Philosophical Logic*, 10, 217–234.

- Lycan, W. G. (2001). *Real conditionals*. Oxford: Oxford University Press.
- MacFarlane, J. (2012). Relativism. In D. Graff Fara & G. Russell (Eds.), *The Routledge companion to the philosophy of language* (pp. 132–142). New York: Routledge.
- Mackie, J. L. (1973). *Truth, probability and paradox: Studies in philosophical logic*. Oxford: Oxford University Press.
- Markovits, H., de Chantal, P.-L., & Brisson, J. (2019). Abstract reasoning and the interpretation of basic conditionals. *Thinking & Reasoning*, 25, 1–13. doi: 10.1080/13546783.2018.1452795
- Marr, D. (1982). *Vision*. Cambridge MA: MIT Press.
- Mates, B. (1953). *Stoic logic*. Berkeley CA: University of California Press.
- McDermott, M. (1996). On the truth conditions of certain “if”-sentences. *Philosophical Review*, 105, 1–37.
- McGee, V. (1985). A counterexample to modus ponens. *Journal of Philosophy*, 82, 462–471.
- Merin, A. (2007). Unconditionals. Unpublished manuscript, available at <http://semanticsarchive.net/Archive/WUwZTk5M/unconditionals.pdf>.
- Mill, J. S. (1843/1872). *A system of logic: Ratiocinative and inductive* (eighth edn.). London: Longmans, Green, Reader, & Dyer.
- Milne, P. (2012). Indicative conditionals, conditional probabilities, and the “defective truth-table”: A request for more experiments. *Thinking & Reasoning*, 18, 196–224.
- Nickerson, R. S. (2016). *Conditional reasoning: The unruly syntactics, semantics, thematics, and pragmatics of “if”*. New York: Oxford University Press.
- Nickerson, R. S., Barch, D. H., & Butler, S. F. (2019). Evaluating conditional arguments with uncertain premises. *Thinking & Reasoning*, 25, 48–71. doi: 10.1080/13546783.2018.1465474
- Oaksford, M. & Chater, N. (2003). Conditional probability and the cognitive science of conditional reasoning. *Mind & Language*, 18, 359–379.
- Oaksford, M. & Chater, N. (2007). *Bayesian rationality: The probabilistic approach to human reasoning*. Oxford: Oxford University Press.
- Oberauer, K., Weidenfeld, A., & Fischer, K. (2007). What makes us believe a conditional? The roles of covariation and causality. *Thinking & Reasoning*, 13, 340–369.
- Over, D. E. & Baratgin, J. (2017). The “defective” truth table: Its past, present, and future. In N. Galbraith, D. E. Over, & E. Lucas (Eds.), *The thinking mind: The use of thinking in everyday life* (pp. 15–28). Hove UK: Psychology Press.
- Over, D. E. & Cruz, N. (2018). Probabilistic accounts of conditional reasoning. In L. J. Ball & V. A. Thompson (Eds.), *International handbook of thinking and reasoning* (pp. 434–450). Hove UK: Psychology Press.
- Over, D. E., Douven, I., & Verbrugge, S. (2013). Scope ambiguities and conditionals. *Thinking & Reasoning*, 19, 284–307.
- Over, D. E. & Evans, J. St. B. T. (2003). The probability of conditionals: The psychological evidence. *Mind & Language*, 18, 340–358.
- Over, D. E., Evans, J. St. B. T., & Elqayam, S. (2010). Conditionals and non-constructive reasoning. In M. Oaksford & N. Chater (Eds.), *Cognition and conditionals: Probability and logic in human thinking* (pp. 135–151). Oxford: Oxford University Press.
- Over, D. E., Hadjichristidis, C., Evans, J. St. B. T., Handley, S. J., & Sloman, S. A. (2007). The probability of causal conditionals. *Cognitive Psychology*, 54, 62–97.
- Pfeifer, N. & Douven, I. (2014). Formal epistemology and the new paradigm psychology of reasoning. *Review of Philosophy and Psychology*, 5, 199–221.
- Pfeifer, N. & Kleiter, G. D. (2010). The conditional in mental probability logic. In M. Oaksford & N. Chater (Eds.), *Cognition and conditionals: Probability and logic in human thinking* (pp. 153–173). Oxford: Oxford University Press.
- Politzer, G., Over, D. E., & Baratgin, J. (2010). Betting on conditionals. *Thinking & Reasoning*, 16, 172–197.

- Ramsey, F. P. (1929/1990). General propositions and causality. In his *Philosophical Papers*, edited by D. H. Mellor (pp. 145–163). Cambridge: Cambridge University Press, 1990.
- Rescher, N. (2007). *Conditionals*. Cambridge MA: MIT Press.
- Rieger, A. (2006). A simple theory of conditionals. *Analysis*, 66, 233–240.
- Rieger, A. (2013). Conditionals are material: The positive arguments. *Synthese*, 190, 3161–3174.
- Rieskamp, J. (2008). The probabilistic nature of preferential choice. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 34, 1446–1465.
- Rieskamp, J. & Otto, P. E. (2006). SSL: A theory of how people learn to select strategies. *Journal of Experimental Psychology: General*, 135, 207–236.
- Rothschild, D. (2014). Capturing the relationship between conditionals and conditional probability with a trivalent semantics. *Journal of Applied Non-classical Logics*, 24, 144–152.
- Sanford, D. H. (1989). *If P, then Q*. London: Routledge.
- Schiffer, S. (2003). *The things we mean*. Oxford: Oxford University Press.
- Self, S. G. & Liang, K. (1987). Asymptotic properties of maximum likelihood estimators and likelihood ratio tests under nonstandard conditions. *Journal of the American Statistical Association*, 82, 605–610.
- Simon, H. A. (1982). *Models of bounded rationality*. Cambridge MA: MIT Press.
- Singmann, H. & Kellen, D. (2013). MPTinR: Analysis of multinomial processing tree models in R. *Behavior Research Methods*, 45, 560–575.
- Singmann, H., Klauer, K. C., & Over, D. E. (2014). New normative standards of conditional reasoning and the dual-source model. *Frontiers in Psychology*, 5, 316.
- Skorupski, J. (1989). *John Stuart Mill*. London: Routledge.
- Skovgaard-Olsen, N. (2016a). Ranking theory and conditional reasoning. *Cognitive Science*, 40, 848–880.
- Skovgaard-Olsen, N. (2016b). Motivating the relevance approach to conditionals. *Mind & Language*, 31, 555–579.
- Skovgaard-Olsen, N., Singmann, H., & Klauer, K. C. (2016). The relevance effect and conditionals. *Cognition*, 150, 26–36.
- Spohn, W. (2013). A ranking-theoretic approach to conditionals. *Cognitive Science*, 37, 1074–1106.
- Spohn, W. (2015). Conditionals: A unifying ranking-theoretic perspective. *Philosophers' Imprint*, 15, 1–30. [hdl.handle.net/2027/spo.3521354.0015.001](https://hdl.handle.net/2027/spo.3521354.0015.001)
- Stalnaker, R. (1968). A theory of conditionals. In N. Rescher (Ed.), *Studies in logical theory* (pp. 98–112). Oxford: Blackwell.
- Stalnaker, R. (1975). Indicative conditionals. *Philosophia*, 5, 269–286.
- Strawson, P. F. (1950). On referring. *Mind*, 59, 320–344.
- van Wijnbergen-Huitink, J., Elqayam, S., & Over, D. E. (2015). The probability of iterated conditionals. *Cognitive Science*, 39, 788–803.
- Vidal, M. & Baratgin, J. (2017). A psychological study of unconnected conditionals. *Journal of Cognitive Psychology*, 29, 769–781.
- Weirich, K. (in press). Defending truth values for indicative conditionals. *Philosophical Studies*.