If p, then p!
Or: A crisis of *Identity*

Matthew Mandelkern*

Draft of July 22, 2019

> All the dialecticians in common say that a conditional is sound when its finisher follows from its leader. But on the question of when it follows, and how, they disagree with one another.
>
> Sextus Empiricus: *AM VIII.112*

Abstract

The *Identity* principle says that conditionals with the form ⌜If p, then p⌝ are logical truths. *Identity* is overwhelmingly plausible, and has rarely been explicitly challenged. But a wide range of conditionals nonetheless invalidate it. I explain the problem, and argue that the culprit is the principle known as *Import-Export*, which we must thus reject. I then explore how we can reject *Import-Export* in a way that still makes sense of the intuitions that support it, arguing that the differences between indicative and subjunctive conditionals play a key role in solving this puzzle.

1 Introduction

Sextus Empiricus's summation in the epigraph remains apt. Clearly a conditional in *some* sense says that the consequent follows from the antecedent; there remains a

* All Souls College, Oxford, OX1 4AL, United Kingdom, matthew.mandelkern@all-souls.ox.ac.uk.
Many thanks to audiences at Oxford, NYU, MCMP, and UCL, and to Andrew Bacon, Kyle Blumberg, Susanne Bobzien, David Boylan, Sam Carter, Ivano Ciardelli, Nilanjan Das, Kevin Dorst, Kit Fine, Branden Fitelson, Vera Flocke, Simon Goldstein, Jeremy Goodman, Ben Holguín, Justin Khoo, Cameron Domenico Kirk-Giannini, Angelika Kratzer, Vann McGee, Eliot Michaelson, Jonathan Phillips, Milo Phillips-Brown, Daniel Rothschild, Ian Rumfitt, Bernhard Salow, Philippe Schlenker, Ginger Schultheis, and Robert Stalnaker for very helpful feedback and discussion. Thanks especially to Cian Dorr for valuable discussion that helped shape key points in the second half of the paper.

great deal of controversy about what kind of following is involved.[1] But plausibly, on any reasonable way of cashing out “following”, any sentence will follow from itself. And so, given any way of making precise this broad way of thinking about conditionals, sentences of the form ⌜If p, then p⌝ should be logical truths.

In the first part of this paper, I show that, despite the overwhelming plausibility of the principle that ⌜If p, then p⌝ is always a logical truth—the *Identity* principle—a wide variety of theories of the conditional invalidate it. I then argue that the culprit behind this failure is the *Import-Export* principle, which says that ⌜If p, then if q, then r⌝ and ⌜If p and q, then r⌝ are always materially equivalent. I show that there is a deep and surprising tension between *Import-Export*, on the one hand, and, on the other, *Identity* (as well as its more general cousin *Logical Implication*, which says that ⌜If p, then q⌝ is always true when p logically entails q). In light of the overwhelming plausibility of the latter principles, we should thus reject *Import-Export*.

In the second part of the paper, I explore how to reject *Import-Export* while still accounting for the intuitive evidence that supports it. Surprisingly, intuitions concerning *Import-Export* seem to diverge for indicatives versus subjunctives: we find concrete counterexamples to *Import-Export* for subjunctives, but apparently not for indicatives. To account for this, I propose a *local* implementation of a widely accepted account of the difference between indicatives and subjunctives, on which indicatives, but not subjunctives, presuppose that the closest antecedent-world is epistemically possible. I show that on the resulting account, *Import-Export* is logically *invalid* for both indicatives and subjunctives—as desired—but it still holds for indicatives in a more limited sense: namely, in all the cases where an indicative’s epistemic presupposition is satisfied. I argue that this is the best way to account for intuitions in this area while still validating *Logical Implication*.

---

1 On the material view, the consequent must follow in the sense of being true at the actual world if the antecedent is (Grice 1989, Lewis 1976, Jackson 1987, Williamson 2019); on the strict view, in the sense of being true at all accessible antecedent worlds (Warmbrod 1981a,b, von Fintel 1999, 2001, Gillies 2007); on the variably strict view, in the sense of being true at all closest antecedent worlds (Stalnaker 1968, Stalnaker & Thomason 1970, Lewis 1973); on probabilistic views, in the sense of being suitably probable on the antecedent (Adams 1975, Edgington 1986).

2 A crisis of *Identity*

Identity, again, says that sentences with the form ⌜If p, then p⌝ are logical truths. *Identity* is a corollary of a more general principle, *Logical Implication*, which says that, when p logically entails q, then ⌜If p, then q⌝ is a logical truth (‘logical entailment’ is, here and throughout, understood in its classical sense: preservation of truth in all intended models). In this section I will focus on *Identity*; later in the paper I will return to the more general *Logical Implication* principle. *Identity* is one of the most natural, and least controversial, principles in the logic of the conditional. Arló-Costa & Egré (2016) call it ‘constitutive of the very notion of conditional’. This seems correct: it seems beyond doubt that, if p holds, then p holds. I will begin by showing that, despite *Identity*’s plausibility—and the lack of explicit challenges to it in the literature—*Identity* is invalidated by a wide range of current theories of the conditional, in particular those (apart from the material conditional) which validate the *Import-Export* principle.

I will henceforth work with a simple propositional language with atoms $A, B, C \ldots$, Boolean connectives ‘$\wedge$’, ‘$\neg$’, ‘$\vee$’; the material conditional ‘$\supset$’ (so $p \supset q$ abbreviates $\neg p \vee q$), and the material biconditional ‘$\equiv$’ (so $p \equiv q$ abbreviates $(p \supset q) \wedge (q \supset p)$). Finally, we have a conditional connective ‘$>$’; later in the paper I will distinguish the indicative conditional connective ‘$>_i$’ from the subjunctive one ‘$>_s$’, but for now I use just one connective ‘$>$’ which ranges over both indicatives and subjunctives. Lower-case italics range over sentences; for readability I use sentences autonymously. Where Γ is a set of sentences of our language, ‘$\Gamma \models p$’ means that Γ logically entails p, in the standard sense that p is true in every intended model where all the elements of Γ are true. We write ‘$p \models q$’ for ‘$\{p\} \models q$’; and ‘$\models p$’ for ‘$\varnothing \models p$’, i.e. to say that p is a logical truth. I emphasize that the use of a formal language is just to facilitate readability, so $p > q$ is just an abbreviation of ⌜If p, then q⌝.[2] With this notation in hand, I restate the three key principles which will play a central role in what follows, using ‘$\rightarrow$’ as a meta-language material conditional:

2 Kratzer (1986) famously argues that it is a mistake to treat ‘if’ as a two-place connective, as I will here. But Khoo (2013) convincingly showed that this question about the syntax of conditionals does not bear on results of the kind I will be discussing here; a parallel argument to Khoo’s shows that my points go through regardless of the syntax of conditionals.

- *Logical Implication (LI)*: $p \models q \rightarrow \models p > q$
- *Identity*: $\models p > p$
- *Import-Export (IE)*: $\models (p > (q > r)) \equiv ((p \wedge q) > r)$

LI and *Identity* are self-explanatory. *IE* is a bit more complicated: it says that what we do with two successive conditional antecedents is the same as what we do with the corresponding conjunctive antecedent. So, for instance, *IE* says that pairs like the following are generally equivalent:

(1) a. If the coin is flipped, then if it lands heads, we'll win.
b. If the coin is flipped and it lands heads, we'll win.

And likewise for the subjunctive version:

(2) a. If the coin had been flipped, then if it had landed heads, we would have won.
b. If the coin had been flipped and it had landed heads, we would have won.

All three of these principles are *prima facie* plausible. The plausibility of *LI* and *Identity* is, I take it, manifest; the plausibility of *IE* comes from the felt equivalence of the pairs in (1) and (2). Later on we will explore in more detail the case for and against each of these principles.

With this background on the table, let me turn to the central claim of this section: that all existing theories of the conditional, apart from the material conditional, which validate *IE* also *in*validate *Identity* (and thus *LI*). To explain this point, start by thinking about what it takes to validate *IE*. *IE* says, in essence, that information in subsequent antecedents is agglomerated: a conditional with two antecedents is evaluated in the same way as a conditional with one corresponding conjunctive antecedent. That means that, to validate *IE*, we need some way of "remembering" successive conditional antecedents. To see why, suppose instead we adopt a classic variably strict view like that of Stalnaker 1968, which does not have a mechanism to do this. On Stalnaker's view, $p > q$ is true just in case q is true at the closest p-world (see §6 for more exposition). So $p > (q > r)$ says that r is true at the closest q-world to the closest p-world. By contrast, $(p \wedge q) > r$ says that r is true at the closest $p \wedge q$-world. A little reflection shows that these truth-conditions are orthogonal, and so *IE* is invalid. What we need to validate *IE*, instead, is some way of keeping track of successive conditional antecedents, and then using these together to evaluate the most deeply-embedded consequent.

Different *IE*-validating theories of the conditional have different mechanisms for doing this. For instance, in McGee (1985)'s framework, conditional antecedents are added sequentially to a set; the consequent is then evaluated at the closest world where all the sentences in that set are true (more or less; see the first appendix for a more careful exposition). In Kratzer (1981, 1991)'s framework, conditional antecedents are similarly added to the value of a "modal base" function which takes each world to a set of propositions, which in turn provides the domain of quantification for evaluating the consequent. Similar approaches are developed in von Fintel 1994, Dekker 1993, Gillies 2004, Starr 2014, Gillies 2009. These theories differ in important ways, but they all have a parameter which is in the business of somehow remembering successive conditional antecedents, so that these can be agglomerated when we arrive at the consequent. Intuitively, that is exactly what is needed in order to validate *IE*: the interpretation of conditionals must depend on a shiftable domain parameter of some kind which gets updated by conditional antecedents, which are then somehow agglomerated.[3]

Structurally, this has an important consequence. What proposition a conditional expresses depends on the setting of this shiftable domain parameter. And thus, since this parameter changes under conditional antecedents, what proposition a conditional expresses can change depending on whether it is embedded under a conditional antecedent. Now consider a sentence with the form $p > p$ and suppose that p itself contains a conditional. Then the first instance of p will be interpreted relative to a different shiftable domain parameter from the second p: when we get to the second (but not the first), that shiftable domain parameter will have been updated with the information that p is true. And that, in turn, means that the two instances of p can

3 As Khoo & Mandelkern (2018), Mandelkern (2018) discuss, not all these systems validate *IE* when what gets imported/exported is itself a conditional. However, that case is not relevant for our purposes, so I will gloss over this detail here.

express different propositions, and so the conditional as a whole can end up being false.

More concretely, think about a conditional of the form $(\neg(A > B) \wedge B) > (\neg(A > B) \wedge B)$, where A and B are arbitrary atoms (in §4.1, we will get even more concrete, looking at conditionals in natural language with this form).[4] This has the form $p > p$. Now consider what happens when we arrive at the consequent of this conditional if we have an *IE*-validating system. At that point, the antecedent will have been added to our shiftable domain parameter. So the shiftable domain parameter will now entail the antecedent, and so in particular will entail B. That means that the parameter will only make available B-worlds for the evaluation of conditionals. So we will end up evaluating the consequent of the conditional at a domain which comprises only B-worlds. The consequent, again, says $(\neg(A > B) \wedge B)$. Focus on the first conjunct, which says a certain conditional, $A > B$, is false. The problem is that if our conditional domain—the domain of worlds which matter for evaluating the conditional—includes only B-worlds, then this conditional—on any reasonable theory of the conditional—can't be false. That means that this conditional, as it appears in the consequent of our target conditional, must be true; and so its negation must be false. So the whole consequent of the conditional will be false, and so the conditional as a whole will be false, provided only that its antecedent is possible (which it often will be).

This gives a sense of why theories which validate *IE* (apart from the material conditional) invalidate *Identity*. In the appendix, I go through this reasoning in more detail in the context of McGee's theory.[5] For now, the crucial point is that theories of the conditional that validate *IE* (apart from the material conditional) dramatically

4 One might worry that conditionals cannot be embedded in the antecedents of conditionals. Indeed, if we restrict our language so that conditionals only have non-conditional antecedents, then we can unproblematically validate *IE* and *LI* together (as in McGee 1989, Arló-Costa 2001, Ciardelli 2019). But this restriction is unmotivated: there is no problem embedding either indicative or subjunctive conditionals in conditional antecedents, as in 'If the vase broke if it was dropped, then it wasn't wrapped in plastic', or 'If the vase would have broken if it had been dropped, then it wasn't wrapped in plastic'. Given that these conditionals make perfectly good sense, we can, and should, ask questions about their logic (see Bacon 2015 for similar arguments).

5 von Fintel (1994)'s version of the restrictor theory introduces more flexibility into that framework, so that *IE* comes out as something like a default inference pattern rather than a strict validity; likewise for Stojnić 2016. Increased flexibility does not, however, improve the situation for these theories: *Identity* still fails for precisely the same reason as on theories which validate *IE* in general.

invalidate *Identity*: they predict that sentences with the form $p > p$ can fail to be true when p contains a conditional, and indeed that such sentences are only ever true when their antecedent is impossible.[6]

3 The culprit

The natural question to ask at this point is whether it is an accident that existing theories of the conditional which validate *IE* invalidate *Identity*. Is there a more reasonable way of validating *IE* that does not lead to failures of *Identity* in general? In this section, I will argue that there is not, because there is a deep tension between *IE* and the more general *LI* principle, of which *Identity* is a corollary. The discussion in the last section already pointed intuitively towards this tension; in this section I will draw out this tension more precisely, showing that, provided we take on board a weak background assumption that seems beyond serious doubt, the material conditional is the only conditional which validates both *IE* and *LI*.

I know of exactly one theory which has been proposed in the literature which validates both *IE* and *LI*: namely, the material conditional. This fact is little solace, however, because there is overwhelming evidence that the natural language conditional 'If…then…' is not the material conditional. This is the consensus view for the subjunctive conditional ('If it had rained, the picnic would have been cancelled'); and it is nearly the consensus view for the indicative conditional ('If it rains, the picnic will be cancelled'). A quick way to see the implausibility of the material analysis is that, since $p \supset q$ is equivalent to $\neg p \vee q$, its negation is equivalent to the conjunction $p \wedge \neg q$. But it is clear that the negation of the natural language conditional ⌜If p, then q⌝ is not equivalent to ⌜p and not q⌝. For instance, 'It's not the case that, if Patch had been a rabbit, she would have been a rodent' and 'It's not the case that, if Patch is a rabbit, she is a rodent' are both clearly true, thanks just to taxonomic facts—whether or not Patch is a rabbit. So neither of these conditionals is equivalent to 'Patch is a rabbit and not a rodent', *pace* the material view.[7] That it

6 That, of course, means that the internal negation of these sentences are logically true: in other words, that $p > \neg p$ can be a logical truth according to these theories, even when p is possible.

7 One might try to appeal to general Gricean considerations to explain divergences in the truth conditions versus assertability conditions of conditionals. But those considerations would not do anything to explain the fact that we *fail* to infer ⌜p and not q⌝ from ⌜Not: if p, then q⌝; Gricean tools are apt for

is possible to validate both *LI* and *IE* by adopting the material analysis is thus cold comfort.

Although the material conditional is the only theory that has been proposed which validates both *LI* and *IE*, it is not the only logically possible one. However, I will argue that there is no *plausible* way to validate *LI* and *IE* together. For, provided we take on a third, very weak background principle—one which seems to me beyond serious doubt—the material conditional is indeed the only connective which validates both *LI* and *IE*. This third principle says that, if $p > q$ and $p > \neg q$ are both true, then p is false. I'll call the principle *Triviality*:

- *Triviality*: $\{p > q, p > \neg q\} \models \neg p$

I'll return in a moment to the motivation for *Triviality*. For now, let me summarize the reasoning which shows that the only connective which validates *LI*, *IE*, and *Triviality* is the material conditional (this reasoning is spelled out in more detail in the second appendix). *LI* says that any conditional with the form of (3) is logically true:

(3) $(\neg(p > q) \wedge q) > \neg(p > q)$

That's because (3) has the form $(a \wedge b) > a$. *IE* and *LI* together entail that the internal negation of this conditional, in (4), is logically true:

(4) $(\neg(p > q) \wedge q) > (p > q)$

That's because (4) is equivalent, by *IE*, to $((\neg(p > q) \wedge q) \wedge p) > q$, which is logically true by *LI*. Thus it follows from *Triviality* that the antecedent of (3) and (4)—namely, $(\neg(p > q) \wedge q)$—is logically false. But, given *Triviality*, if $(\neg(p > q) \wedge q)$ is logically false, then our conditional must validate *Modus Ponens* (*MP*), which says from p and $p > q$, we can infer q. And it is a well-known fact, due to Gibbard (1981), that if *IE*, *LI*, and *MP* are all valid, then the conditional must be the material conditional.

explaining how inferences get amplified, but not very useful for explaining how logical entailments get blocked.

In sum: any conditional which validates *IE*, *LI*, and *Triviality* has to be the material conditional. Since 'If. . . then. . . ' is not the material conditional, these three principles cannot all be valid for the natural language conditional 'If. . . then. . . '.

Before discussing how to respond to this result, let me briefly situate it in relation to Gibbard (1981)'s famous result, which, again, showed that *LI*, *IE*, and *MP* can be jointly validated only by the material conditional. I have shown that the same result follows if we replace *MP* with *Triviality*. And *Triviality* is strictly weaker than *MP*: any theory of the conditional which validates *MP* validates *Triviality*, but not *vice versa*. So the present result strengthens Gibbard's by replacing *MP* with a much weaker principle. I will presently argue, moreover, that *Triviality* has a much firmer dialectical status than *MP*. If that is right, then the present strengthening of Gibbard's result shows that there is a tension that has been missed in the response to his result: the fundamental tension underlying his result is not between *IE* and *MP*, as many have thought, but is rather between *IE* and *LI*.

4 Responses

We have identified a tension between three principles. Which one should we reject?[8]

4.1 *Triviality*

Consider first *Triviality*. The most direct evidence for *Triviality* comes from logical and mathematical contexts. In such contexts, a very natural way to argue that p is false is to show that, if p is true, then q is true, and if p is true, then q is false; from which we can conclude that p is false. This reasoning, however, is only valid if *Triviality* is.

8 I should note that all three principles might fail once we admit semantic vocabulary ('true', 'false', etc.) into our language, due to semantic paradoxes (see e.g. Field 2008). Having said that, I think that there is much to be gained by focusing initially on the logic for a fragment free of semantic vocabulary, in the hopes that a theory that incorporates semantic vocabulary will be able to build directly on this (see Field 2016 for an approach to the conditional that takes exactly this approach). Thanks to Cian Dorr for helpful discussion.

Another way to motivate *Triviality* is by way of two more general principles which are almost universally accepted, and which entail *Triviality*.[9] The first is the *Agglomeration* principle, which says that $p > q$ and $p > r$ together entail $p > (q \wedge r)$. *Agglomeration* is, in Hawthorne (2005)'s words, 'overwhelmingly intuitive': if something would obtain if p does, and some other thing would obtain if p does, then surely both things would obtain if p does (e.g., we can infer 'If it rains, the picnic will be cancelled and the parade will be cancelled' from 'If it rains, the picnic will be cancelled' and 'If it rains, the parade will be cancelled'). The second principle says that, if $p > \bot$ is true, then p is false, where $\bot$ is any contradiction. This principle is, again, very compelling: if $p > \bot$ is true, then $\bot$ must in some sense follow from p; but $\bot$ cannot follow in any relevant sense from a truth. Again, this principle is applied most often in logical contexts, where showing that if p holds, then some contradiction holds, is taken as conclusive evidence that p does not hold. These two principles together entail *Triviality*.

Finally, *Triviality* is an immediate consequence of *Weak Conditional Non-Contradiction*, which says that, whenever $\Diamond p$ is true, $p > q$ and $p > \neg q$ cannot both be true. *Weak Conditional Non-Contradiction* is almost universally accepted; assuming that the relevant notion of possibility is reflexive, it entails *Triviality*.

Thus it is very hard to reject *Triviality*. Again, nearly every theory of the conditional validates it, even those theories which reject *MP*.[10] And indeed, while there has been a serious case made against *MP* by McGee 1985, there has been no serious case made against *Triviality*; and I do not see anyway to convert the standard case against *MP*, based on complex conditional consequents, into a case against *Triviality*. So I do not think there is any motivation for rejecting *Triviality*.

9 The only theories I know of which invalidates *Agglomeration* (and so invalidate *Triviality*) treat the conditional as an *existential* operator, so that $p > q$ says *some* accessible p-world is a q-world (Bassi & Bar-Lev 2018, Herburger 2018).

10 Apart from the existential theories mentioned above, the only other theory I know of that invalidates *Triviality*, as Ivano Ciardelli has pointed out to me, is the theory that comprises the conditional from Kolodny & MacFarlane 2010, together with the semantics for epistemic modals from Yalcin 2007, MacFarlane 2011. On that theory, $p > q$ says that all maximal p-accepting substates of the global information state s accept q. Given Yalcin's semantics for epistemic modals, $(\Diamond p \wedge \neg p) > q$ and $(\Diamond p \wedge \neg p) > \neg q$ are logical truths for any q, but $\neg(\Diamond p \wedge \neg p)$ is not a logical truth. This is interesting, but I think largely orthogonal to present issues: the way in which this theory of the conditional avoids the present results is by invalidating *IE*.

4.2 *Logical Implication*

So let us explore instead the possibility of rejecting *LI* or *IE*. Unlike rejecting *Triviality*, both of these options have been taken seriously in the literature. Although almost no one has to my knowledge explicitly advocated rejecting *LI*, a broad range of theories *do* reject it, as we have just seen; and it is of course well known that many theories of the conditional invalidate *IE*.

Is there a serious case against *LI*? Perhaps against parts of it. In particular, one consequence of *LI* is that all conditionals of the form $\bot > p$ are true, since $\bot$ classically entails everything. This corollary is widely but not universally accepted: there is a serious case for rejecting it (see e.g. Jenny 2016 and citations therein). However, this instance of *LI* is the only one I know of which has been explicitly challenged in the literature.[11] But this particular instance plays no role in the proof of the collapse result given in the last section. Indeed, that result depends only on three instances of *LI*: namely, the claim that $(p \wedge q) > p$ is a logical truth, that $((p \supset q) \wedge p) > q$ is a logical truth, and that $p > q$ is a logical truth whenever q is. It is hard to see how these mild instances of *LI* can be rejected. Even if we don't want to accept that $\bot > p$ is always a logical truth, I can't see any serious case against a more limited form of *LI*, restricted to conditionals with consistent antecedents—and that is all we need to obtain the result presented in the last section. Indeed, we could focus, even more narrowly, on *Identity*, since, as we have seen, even that very simple corollary of *LI* is invalidated by all extant non-material theories which validate *IE*. It is very hard to see any case against a principle like *Identity*.[12]

---

11 With the exception of Sextus Empiricus's famously obscure 'emphasis' account, which invalidates *Identity* because 'it is impossible that anything be included in itself' (*PH 2.112*).

12 It is important to distinguish *LI* from some related principles which appear similar but are much less defensible. First, consider *Multi-Premise LI (MPLI)*, which says that, if $\Gamma, p \models q$, then $\Gamma \models p > q$. *MPLI* is definitely false, for from *MPLI* alone we can derive the falsehood that $p \supset q \models p > q$ (Bonevac et al. 2013). Despite its superficial similarity to *LI*, however, it is easy to find intuitive grounds for rejecting *MPLI*: on any plausible theory, $p > q$ asks us to evaluate q at a range of *potentially non-actual* worlds. p will plausibly hold at all those worlds, but other things that are true at the actual world may not hold. In particular, then, all the elements of Γ may hold at the actual world but fail to hold at some of the relevant p-worlds; in which case the fact that Γ and p together entail q does not do anything to help make the conditional $p > q$ true at a given world, even if Γ happens to be true there. (More concretely: suppose the coin landed heads. It doesn't follow that, if the coin had landed tails, then it would have landed heads—because the antecedent precisely takes us to a world where we do not hold fixed the proposition that it landed heads.) So there are very

Even in the absence of any theoretical case against *LI*, we should still explore natural language conditionals where it might plausibly fail. And indeed, we have already precisely pinpointed the form of conditionals which *LI* says should be logically true, and which *IE*-validating theories by contrast say should not be: *LI* says that sentences with the form $(\neg(p > q) \wedge q) > (\neg(p > q) \wedge q)$ and $(\neg(p > q) \wedge q) > \neg(p > q)$ will invariably be true, while *IE* says they will invariably be false, provided the antecedents are consistent. The same points go for conditionals which reverse the order of the conjuncts, which seem easier to process, so I will focus on sentences with that form. Consider first the conditionals in (5):

(5) a. If the vase had broken, but it's not the case that the vase would have broken if it had been wrapped in plastic, then it's not the case that the vase would have broken if it had been wrapped in plastic.

b. If the match had lit, but it's not the case that the match would have lit if it had wet, then it's not the case that the match would have lit if it had been wet.

These have the form $(q \wedge \neg(p > q)) > \neg(p > q)$, so *LI* predicts they are trivially true. By contrast, *IE*-validating theories predict that, provided the antecedents are possible, these are necessarily false. It seems easy to imagine cases where these antecedents are possible; in those cases, *IE*-validating theories predict these conditionals are necessarily false. That prediction, however, seems wrong: these conditionals sound logically true. Roughly speaking, the antecedent of (5a) seems to imagine a situation where two things are true: namely, that the vase broke; but that plastic wrapping would have prevented this. In such a situation, the vase broke; and plastic wrapping

---

natural reasons to reject *MPLI*; but these are not also reasons to reject *LI*, since if *p alone* entails q, then any range of relevant p-worlds will all be q-worlds. Second, consider *Converse LI (CLI)*, which says that, if $\models p > q$, then $p \models q$. I am inclined to accept *CLI*, and it is valid on the theory I adopt below. But its dialectical status is insecure, since it is in obvious tension with *IE*: for instance, *IE* says that $p > (q > p)$ is a logical truth, but of course proponents of *IE* will not want to say that $p \models q > p$. Insofar as we take the intuitive evidence for *IE* as decisive, then, we will not be at all inclined to accept *CLI*, and we will again have a good explanation of why it fails: because conditional antecedents agglomerate in a way that a conditional antecedent together with a logical premise in an argument do not. But, once more, this reasoning does not give us good grounds for rejecting *LI* (let me emphasize, again, that I do not accept this reasoning, and that I do accept *CLI*; my point here is just that *LI* is on a firmer dialectical foundation than *CLI* or *MPLI*).

would have prevented this; so the consequent must be true in such a situation; so the whole conditional must be true in any situation. Similar reasoning suggests that (5b) is logically true as well.

Conversely, *IE*-validating theories predict that the internal negations of the conditionals in (5) are logical truths, as in (6):

(6) a. If the vase had broken, but it's not the case that the vase would have broken if it had been wrapped in plastic, then the vase would have broken if it had been wrapped in plastic.

b. If the match had lit, but it's not the case that the match would have lit if it had wet, then the match would have lit if it had been wet.

Again, this seems wrong: these conditionals sound trivially false to me, and similar reasoning to that just rehearsed seems to bear this out. So it looks like the predictions of *LI*, not *IE*, are correct. The same goes for nearby variants, like conditionals with the form $(q \wedge \neg(p > q)) > (q \wedge \neg(p > q))$.[13]

Matters are similar for indicatives. Consider for instance the indicative versions of (5b) and (6b), in (7):

(7) a. If the match lit, but it's not the case that it lit if it was wet, then it's not the case that it lit if it was wet.

b. If the match lit, but it's not the case that it lit if it was wet, then it lit if it was wet.

Again, *LI* predicts that (7a) is a logical truth, and (7b) false whenever its antecedent is possible; *IE*-validating theories predict the reverse. Once more, the predictions of *LI* seem correct. (There is an interesting further element here, which is that both the conditionals in (7) strike me as somewhat odd; I return to this point in §7.4.)

13 It is crucial throughout that context-sensitive language like pronouns, implicit temporal and locative indexing, and so on remain fixed throughout the instances of p and q; obviously things are different if they do not. Readers worried about these issues can add overt temporal and locative phrases to these sentences and throughout to control for this; as long as this is done consistently, it does not change judgments about the sentences.

Intuitions about natural language thus do not give us any grounds for rejecting *LI*: as far as I can tell, instances of *LI* sound just as valid as we might pretheoretically expect them to; by contrast, the predictions of *IE* in the same cases seem wrong.

These intuitions also give us reason to reject a different kind of response to our result. One could adopt a non-classical, "dynamic" notion of entailment, and argue that validating *LI*, *Triviality*, and *IE* as interpreted using this dynamic notion suffices to account for intuitions in this area.[14] The problem with this is that dynamic notions of entailment diverge profoundly from classical ones in the central cases that concern us here: in particular, neither p nor $p \wedge q$ dynamically entails p on standard dynamic notions.[15] So, even if we validate a dynamic version of *LI*, we still invalidate principles like *Identity*, whether understood dynamically or statically. Indeed, this is exactly the situation for the standard dynamic conditional from Dekker 1993, Gillies 2004,[16] as well as for the different dynamic theory in Gillies 2009. But as we have just seen, sentences with the form $p > p$ and $(p \wedge q) > p$ *do* seem to be logically valid. So moving to a dynamic version of *LI* won't account for the intuitions that motivate *LI* in the first place. In other words, in thinking about how to respond to the present result, we can largely bracket abstract questions about how best to think about logical entailment in general and focus on questions like: are sentences of the form $(\neg(p > q) \wedge q) > (\neg(p > q) \wedge q)$ logical truths? Both theoretical and intuitive evidence suggest that they are; but dynamic theories, like our other targets here, do not have the resources to account for this, since even if they validate a dynamic variant of *LI*, they still invalidate *Identity*.

4.3 *Import-Export*

If we want to hold onto both *Triviality* and *LI* and we do not want to collapse to the material conditional, then we must reject *IE*. And indeed, I think that is precisely what we should do: in light of the tension between these three principles, and the evidence for the first two, we have good evidence that *IE* must be invalid.

14 Following Veltman 1996; compare Gillies (2009), Bledin (2015)'s responses to Gibbard's result and to McGee's putative counterexamples to *MP*, respectively.

15 E.g. on the first two notions of dynamic entailment from Veltman 1996.

16 On that theory, $s[p > q] = \{w \in s : s[p] = s[p][q]\}$. So for instance on that theory, $(\neg(p > q) \wedge q) > (\neg(p > q) \wedge q)$ will not be everywhere accepted.

But these considerations are rather indirect; it would be nice to either find more direct evidence against *IE*, or else find a reasonable explanation for the lack of such evidence. In the rest of this paper, I will do both these things. For it turns out that subjunctive and indicative conditionals behave very differently with respect to *IE*: subjunctives yield intuitive counterexamples to *IE*, while indicatives appear not to.

There is, again, a good case to be made for *IE*. The central evidence for *IE* comes from the felt equivalence of pairs like those in (1) and (2). Many other similar pairs have been given in the literature, and they do tend to feel pairwise equivalent; see van Wijnbergen-Huitink et al. (2014) for experimental evidence that confirms these intuitions. Another, more abstract, motivation for *IE* comes from Ramsey (1978/1931)'s famous suggestion that you should believe $p > q$ iff you believe q after adding p hypothetically to your stock of beliefs. Repeated application of this test suggests that you should believe $p > (q > r)$ iff you should believe $(p \wedge q) > r$—which in turn suggests that these have the same truth-conditions (see Arló-Costa 2001).

In the case of subjunctive conditionals, however, *IE* seems to break down. For instance, (8) (from Etlin 2008) and (9) (from Stephen Yablo, p.c.) instantiate the *IE* schema, but are not intuitively equivalent:

(8) a. If the match had lit, then it would have lit if it had been wet.
 b. If the match had lit and it had been wet, then it would have lit.

(9) a. If I had been exactly 6 feet tall, then if I had been a bit taller than 6 feet, I would have been 6'1".
 b. If I had been exactly 6 feet tall and a bit taller than 6 feet, I would have been 6'1".

(8a) can intuitively be false in some circumstances, whereas (8b) intuitively cannot. Conversely, (9a) is plausibly true in some, but not all, circumstances; whereas (9b) feels very different (if it is ever true, then it is surely only true trivially, unlike (9a)). These felt inequivalences target the two directions of *IE*, and suggest that neither direction is valid in general for subjunctives. Importantly, these cases are not outliers: it is straightforward to generate further inequivalent pairs like this on a similar model, as in (10):

(10) a. If the exams had been marked, then if the faculty had gone on strike, then the exams would still have been marked.

b. If the exams had been marked and the faculty had gone on strike, then the exams would still have been marked.

(10b) is obviously true, whereas we can certainly imagine (10a) being false, if the strike would have prevented the exams being marked. And so, as desired, we have direct evidence which matches our indirect evidence: *IE* is not valid for subjunctives.

Things are more complicated, however, for indicative conditionals. Consider the indicative versions of the three pairs we have just looked at:

(11) a. If the match lit, then it lit if it was wet.

b. If the match lit and it was wet, then it lit.

(12) a. If I am exactly 6 feet tall, then if I am a bit taller than 6 feet, then I am 6'1".

b. If I am exactly 6 feet tall and a bit taller than 6 feet, then I am 6'1".

(13) a. If the exams were marked, then if the faculty went on strike, then the exams were still marked.

b. If the exams were marked and the faculty went on strike, then the exams were still marked.

Unlike the corresponding subjunctive pairs, these indicative versions appear to be pairwise equivalent. *IE* is a universal claim, and so the lack of a counterexample in these cases of course does not necessarily show that it is valid for indicatives. But since the very same pairs in the subjunctive mood strike us as inequivalent, it at least suggests that we will not find direct evidence against *IE* for indicatives. And this poses a real puzzle. We have powerful, but indirect, evidence that *IE* is not valid for indicative conditionals, since *Triviality* and *LI* appear to be valid for the indicative, and the indicative is not the material conditional. But we don't seem to find concrete counter-instances to *IE* for indicatives. The goal of the rest of the paper will be to make sense of this puzzling situation.

5 Strawson concepts

What can we say when we have indirect evidence that a principle is not logically valid, but we don't seem to find concrete counter-instances to it? The key idea behind the solution that I will propose here is that an inference pattern might fail to be logically valid, while still having properties that make it hard to find concrete counter-instances to it. In particular, the inference pattern may fail to preserve truth in *all* intended models, but still preserves truth in all cases *where we might plausibly use the sentences in question*. This is roughly the notion that von Fintel (1999) spelled out as *Strawson validity*, following Strawson 1952. Here is my take on the notion (which I should note differs from von Fintel's both in some of its details and its motivations):[17]

> **Strawson entailment**: Γ Strawson entails p iff for any context c and world $w \in c$, if the presuppositions of all the elements of Γ and of p are satisfied in $\langle c, w \rangle$, then if all the elements of Γ are true at $\langle c, w \rangle$, so is p.

If an inference is Strawson valid, it does not necessarily preserve truth in all worlds, but it does preserve truth in any context-world pair where all the premises and the conclusion have their presuppositions satisfied—which includes all contexts where the sentences in question can be naturally used. So, if an inference is Strawson valid, it will be hard to find natural counterexamples to it—even if it is not logically valid.

An example of Strawson entailment comes from agreement features on pronouns, which are plausibly a certain kind of presupposition—in particular *expressive presuppositions* in the idiom of Soames 1989 (see Sudo 2012 for extensive discussion). Consider the sentence 'She is female'. Plausibly, this sentence will be true on any natural occasion where it is used, since 'she' will naturally be used only to talk about a female. There is thus *some* sense in which this sentence is valid. This sentence may not be *logically valid*, though: intuitively, the sentence may fail to be true if 'she' is used mistakenly or non-canonically to denote a female ('I heard Sue got a cat.' 'Yeah, she's a female.' 'No, he's a male.') So the validity of this sentence is

17 The main difference is the relativization of the notion to a context, which plays a crucial role in formulating some presuppositional constraints, including the one I discuss below.

well modeled as a Strawson validity: true on any ordinary occasion of use, even if not always true.

Strawson entailment is in fact just one of a class of what we might call *Strawson concepts*: concepts that stipulate that certain relations hold between sentences *whenever their presuppositions are all satisfied*. So, for instance, we can also spell out a notion of *Strawson acceptance equivalence*, which says that, in any context where the presuppositions of p and q are satisfied, p is accepted throughout that context iff q is.[18] Likewise, we can spell out a corresponding *probabilistic Strawson concept*, which says that p and q invariably have the same probability in any context where their presuppositions are all satisfied, according to that context's associated probability measure.

With these principles in hand, I can state my proposal: I will aim to use Strawson concepts to account for the felt validity of *IE* for indicative conditionals—without actually logically validating *IE*, and thus avoiding the tension between *LI* and *IE*.

Before laying out my proposal, let me briefly say a bit more about Strawson concepts. First: the notion of *presupposition* here is meant to capture a sentence's semantically encoded felicity requirements. This includes semantic presupposition, as well as expressive presuppositions. Here, I will focus exclusively on the latter. Second: a context is a model of a conversation's common ground, Stalnaker 1974b, 1978, which represents in some sense the epistemically accessible worlds in the conversation (or doxastically accessible worlds; for our purposes, it is not important whether the relevant notion is factive or not, though for simplicity I will continue to talk about "epistemic accessibility").

Finally, there are a variety of ways we can formally model Strawson concepts. On the standard trivalent approach, sentences have one of three truth values (0, 1, and #), and p Strawson entails q just in case, whenever p has truth value 1, q has truth value 1 or #. This approach, however, obscures some interesting distinctions. For instance, suppose that we have two sentences, x and y, each of which is true whenever its presuppositions are satisfied. But, while x's truth conditions on their own would (abstracting away from its presuppositions) suffice to guarantee its truth, the truth

18 More formally: if the presuppositions of both p and q are satisfied throughout $\{\langle c,w\rangle : w \in c\}$, then p is true throughout $\{\langle c,w\rangle : w \in c\}$ iff q is.

conditions of y on their own would allow it to be false, even though, whenever it would be false, its presuppositions are not satisfied. For example, x could be 'Either she is a cat or she is not a cat', and y could be 'She is female'. Plausibly, these have the same status from a trivalent point of view: both are always either 1 (when their presuppositions are satisfied) or # (when not, i.e. when 'she' is used to refer to a male). But intuitively these have a different status: the first sentence always *would* be true, just in virtue of its form, if its presupposition were satisfied; whereas the same is not true of the latter. A trivalent framework misses distinctions like this. We can better capture them in a multi-dimensional approach (Herzberger 1973, Karttunen & Peters 1979, Dekker 2008, Sudo 2012).[19] This approach distinguishes different dimensions of content: in particular, truth and expressive presupposition.[20] The first dimension records whether presuppositions are satisfied; the second records truth (as a mnemonic device, I'll underline the presuppositional dimension). Then where $*$ is shorthand for either 1 or 0, we can say, for instance, that p is *Strawson* valid iff $[\![p]\!]^{c,w}$ is always $\langle \underline{1}, 1 \rangle$ or $\langle \underline{0}, * \rangle$; whereas p is *logically* valid iff $[\![p]\!]^{c,w}$ is always $\langle \underline{*}, 1 \rangle$. This lets us distinguish between sentences like x and y above: both are Strawson valid, but only x is also logically valid. Multi-dimensional approaches thus allow us to pull apart the logical features of a system which are due to its truth-conditions from those which are due to its presuppositions, and so I will work with a multi-dimensional system here. I should emphasize, however, that this multi-dimensional setting is separable from my main points, which could easily be recast in a trivalent setting (just read $\langle \underline{0}, * \rangle$ as #, $\langle \underline{1}, 1 \rangle$ as 1, and $\langle \underline{1}, 0 \rangle$ as 0).[21]

19 Thanks to Cian Dior for helpful discussion on this point.

20 Again, I ignore semantic presupposition here. Expressive presuppositions project universally, i.e. throughout the local context; I will leave the projection rules for expressive presuppositions implicit.

21 It is controversial exactly what Strawson concepts tell us. Dorr & Hawthorne (2018) in particular spell out a range of worries about using Strawson entailment in particular to account for the felt validity of inference patterns in natural language. I think that many of those worries are compelling. But I will be putting Strawson entailment to a very limited use here, which I think avoids these worries, namely, to account for the felt validity of a single sentence schema, rather than the validity of an inference from premises to a conclusion. I will give a theory of the conditional on which sentences which instantiate *IE* are always true provided a selection function is chosen in such a way that satisfies the conditionals' presuppositions; since there will in any context be a fair amount of freedom in choosing a selection function, the idea is that we will try to choose one that makes the sentence's presuppositions satisfied. I believe that this limited application of Strawson validity avoids some issues that arise for broader applications of the notion (for instance, the failures of transitivity characteristic to Strawson validity are not relevant for sentence schemas like *IE*).

6 Indicatives versus subjunctives

My proposal, again, is to account for the apparent validity of *IE* for indicatives using Strawson concepts. In the rest of this paper, I will argue that a particular account along these lines in fact falls out of an independently motivated account of more general differences between indicatives and subjunctives. In this section I explore those differences and develop an account of them. Then I will show how this general account helps with *IE*.

Following much of the literature on conditionals, going back to Stalnaker (1975)'s seminal paper, I will assume that indicative and subjunctive conditionals have structurally identical truth conditions, differing only in (i) what selection function we use to evaluate them, and (relatedly) (ii) in their presuppositions. It is to some degree immaterial for our purposes what truth conditions we adopt as our baseline theory, provided that the theory logically validates *LI* and *Triviality*, and logically invalidates *IE*. For concreteness, I will adopt the semantics from Stalnaker 1968, Stalnaker & Thomason 1970, on which a conditional says that the closest antecedent-world is a consequent-world. In more detail, we assume that context provides an indicative selection function f_i and a subjunctive selection function f_s. Selection functions take a proposition p and world w to the closest world to w where p is true.[22] Thus where $>_i$ is the indicative conditional, $p >_i q$ is true at $\langle c, f_i, f_s, w \rangle$ iff q is true at $\langle c, f_i, f_s, f_i(p,w) \rangle$; likewise, where $>_s$ is the subjunctive conditional, $p >_s q$ is true at $\langle c, f_i, f_s, w \rangle$ iff q is true at $\langle c, f_i, f_s, f_s(p,w) \rangle$. Any selection function must satisfy the following conditions:[23]

- *Strong Centering*: $f(p,w) = w$ iff $w \in p$;
- *Success*: $f(p,w) \in p$ provided $p \neq \varnothing$;

22 For readability, I will use italics for both sentences and the corresponding propositions, ignoring relativization to contexts for simplicity. I will sometimes still use Roman letters to range over sentences of natural language, rather than our formal language. Stalnaker relativizes *Non-triviality* to an accessibility relation, but the role of an accessibility relation will be largely played by the compatibility constraint we take on board here, so I will ignore it; my proposal could be recast *via* an accessibility relation.

23 These conditions ensure that each selection function determines a well-order on worlds, relative to each world, which allows us to move freely between talk of a selected p-world from w and the closest p-world to w.

- *CSO*: if $f(p,w) \in q$ and $f(q,w) \in p$, then $f(p,w) = f(q,w)$; and
- *Non-Triviality*: $f(\varnothing,w) = \lambda$, where λ is an absurd world that makes all sentences true.

How can we build on these truth conditions to account for the general differences between indicative and subjunctive conditionals? There are two main proposals in the literature. The first, from Stalnaker 1975, says that indicative conditionals are always evaluated relative to a selection function which treats contextually possible worlds as being closer to each other than any other worlds: i.e., for any conditional antecedent p and world $w \in c$, $f_i(p,w)$ is in c if there is a p-world in c. The motivation for this constraint comes from the observation that, when we leave open that $\neg p$ and we accept $p \vee q$, we generally also accept the indicative conditional $\neg p >_i q$, but not necessarily the corresponding subjunctive $\neg p >_s q$. So, for instance, once we accept (14a), then it seems we must also accept the indicative conditional (14b), but not necessarily the subjunctive (14c):

(14) a. It was the gardener or the butler, and it might have been either.
b. $\rightsquigarrow$ If it wasn't the gardener, it was the butler.
c. $\not\rightsquigarrow$ If it hadn't been the gardener, it would have been the butler.

Stalnaker's constraint is exactly what is needed, in the context of his theory, to account for this inference pattern.

The second proposal, from von Fintel 1998, says that indicatives presuppose that their antecedents are epistemically possible:[24] $p >_i q$ is felicitous only in a context compatible with p, whereas $p >_s q$ can be felicitous even in a context incompatible with p. The motivation for this is the simple observation that, once p is accepted, $\neg p >_i q$ becomes quite weird, while $\neg p >_s q$ remains fine, as illustrated by (15):

(15) John didn't come to the party.
a. #If he came, it was a disaster.

24 See also Gillies 2009, Holguín 2018 for more recent motivation. Dorst (2019) argues against this compatibility constraint. His data, however, seem to me to be order sensitive in a way that suggests that they involve a context shift; there is not adequate space here for careful discussion, but see Holguín (2018) for careful discussion to this effect.

b. If he had come, it would have been a disaster.

Both of these constraints—which are independent from each other—strikes me as being well-motivated, and so, following von Fintel 1998, I conclude that we should take on board both.[25] Indeed, there is a simple way to capture both constraints together in Stalnaker's framework: we say that indicatives presuppose that the selected world is epistemically possible, i.e. $p >_i q$ presupposes $\forall w \in c : f_i(p,w) \in c$. I will call this the *indicative constraint*. This is equivalent to the conjunction of Stalnaker's closeness constraint and von Fintel's compatibility constraint.[26]

This is the fairly standard story which emerges from von Fintel 1998. The only addition I want to propose to this picture is the observation that the indicative presupposition needs to be implemented in a *local* way.[27] To see this point, consider the contrast in (16):

(16) I don't know whether Bob came to the party.

a. #But suppose that Bob came to the party, and that if he didn't come, he went to work.

b. But suppose that Bob came to the party, and that if he hadn't come, he would have gone to work.

The embedded indicative conditional in (16a) is infelicitous, in contrast to the subjunctive variant in (16b). But this is surprising from the point of view of the standard picture just sketched, because, relative to the global context in (16), it is epistemically possible that Bob went to the party, and so, globally speaking, the compatibility part of the indicative constraint seems to be satisfied. To account for the contrast in (16), it looks like we need to compute the compatibility requirement relative to the *local* context which takes into account the information in the left conjunct in (16a)—that Bob came to the party.

25 See Harper (1976), Bacon (2015), Dorr & Hawthorne (2018) for the same (or similar) conclusions.

26 Stalnaker's constraint immediately follows. The compatibility constraint does as well, since $f_i(p,w)$ must be in p provided that $p \neq \varnothing$, in which case it will be λ; but since λ is never in the context, we know that $f_i(p,w) \neq \lambda$, and so $f_i(p,w) \in p$; since $f_i(p,w)$ is epistemically accessible, there must be an epistemically accessible p-world. Conversely, the two constraints together guarantee that f_i must select an epistemically possible world: Stalnaker's constraint guarantees that an epistemically possible world is selected if possible, and the compatibility constraint ensures that this will be possible.

27 Thanks to Irene Heim, Cian Dorr, and Ginger Schultheis for very helpful discussion on these points.

Similar points can be made in a variety of other environments. For instance, consider the pair of quantified conditionals in (17):

(17) I don't know which students studied.

a. #But every student who didn't study passed if she studied.

b. But every student who didn't study would have passed if she had studied.

Again, in the global context of (17), for every student x, we leave it open that x studied. Yet the indicative conditional in (17a), with the antecedent 'if she studied', seems unacceptable in the context of a quantifier whose restrictor is 'didn't study'. By contrast, the subjunctive variant in (17b) is fine. And so again, it looks like the compatibility constraint in question must be calculated relative to a local context which entails the restrictor—'didn't study'—rather than just relative to the global context. Another motivation for a local version of the indicative constraint (which will be especially pertinent to our broader interests here) comes from nested conditionals. Suppose we have a die which is either weighted towards the evens or the odds; we don't know which. Compare (18a) and (18b):

(18) We don't know whether the die was thrown.

a. #If the die was thrown and landed four, then if it didn't land four, it landed two or six.

b. If the die had been thrown and landed four, then if it hadn't landed four it would have landed two or six.

Again, the antecedent of the embedded conditional—that the die didn't land four—is compatible with the global context. But, embedded under a conditional antecedent that entails that the die landed four, only the subjunctive variant in (18b) seems acceptable, while the indicative variant is not. (18b) is perhaps a bit jarring, but it has a clear meaning: it communicates that, if the die had landed four, then it would have been weighted towards evens, and so would have landed two or six if not four. By contrast, the indicative variant in (18a) sounds awful. Once more, it looks like the indicative's compatibility constraint in the consequent of a conditional is calculated relative to a local context: in this case, one which entails the information in the conditional's antecedent.

Conversely, as Kyle Blumberg has pointed out (p.c.), conditionals can be felicitous even when their antecedent has been ruled out in the global context, provided that it remains locally possible, as in (19):

(19) Ann didn't come to the party. But Bill thinks that Ann might have come to the party, and he thinks that if she came to the party, she avoided him.

Again, this is explained if the indicative constraint is calculated relative to the local context—Bill's belief worlds—rather than the global context.

That the indicative constraint is somehow calculated locally in fact looks unsurprising from the point of view of the recent literature on epistemic modality. That literature has suggested that epistemic accessibility is typically calculated in a local manner in general.[28] This can be illustrated by variants on (16)–(18) which replace the indicative ⌜If p. . . ⌝ with ⌜Might p⌝:

(20) a. #I don't know whether Bob came to the party. But suppose that Bob came to the party, and that he might not have.
b. #We don't know which students studied. But every student who didn't study might have studied.
c. #We don't know whether the die was thrown. But if the die was thrown and landed four, then it might not have landed four.

In each of these cases the prejacent of the 'might' is compatible with what is globally epistemically possible; nonetheless the sentences are infelicitous (when the 'might' is read epistemically). We might expect a sentence like (20a) to mean the same as 'Suppose that Bob came to the party, and that, for all we know, he didn't come', which is perfectly coherent; apparently, these differ in meaning. Cases like these suggest that epistemic possibility in general is calculated locally—i.e., in a way which takes into account local information, like the information from a left conjunct, the restrictor of a conditional, or a conditional antecedent. Given all this, it is not

28 See e.g. Kratzer 1981, 1986, Groenendijk et al. 1996, Aloni 2001, Yalcin 2007, Dorr & Hawthorne 2013, Silk 2017, Stojnić 2016, Ninan 2018, Mandelkern 2019.

surprising that the indicative constraint, which connects indicative conditionals to epistemic possibility, should also be calculated locally.

There are different ways we could capture the locality of the indicative constraint. For each "local" theory of epistemic modality, we could build a corresponding local indicative constraint in roughly similar fashion: for instance, we could do so based on the dynamic theory of Groenendijk et al. 1996, Aloni 2001; the domain theory of Yalcin 2007, MacFarlane 2011, Klinedinst & Rothschild 2012; the salience-based theories of Dorr & Hawthorne 2013, Silk 2017, Stojnić 2016; the bounded theory of Mandelkern 2019; or the counterpart theory of Ninan 2018. Here I will build loosely on the bounded theory.

I will not compare this approach in detail to alternative possibilities that we might obtain by taking one of these other approaches, partly for reasons of space and partly because my main aim here is to lay out one possible positive proposal, not to argue that it is the only possible one. Let me briefly note, however, the reason I am not building on a dynamic theory, which would in some way be the most conservative option for localizing the indicative constraint. The issue is that, as I noted above, the dynamic theory invalidates *Identity* and so, I have argued, is not a viable baseline theory of the conditional. In essence, this is because, on dynamic theories (and the other theories which are our targets here), local information shifts the interpretation of embedded conditionals, leading to failures of principles like *Identity*. By contrast, in the theory I develop here, local information does not *shift* the interpretation of embedded conditionals, but rather *constrains* it by way of an expressive presupposition.

To develop this idea, let me briefly say a bit more about the bounded theory. That theory borrows the notion of a *local context* from the theory of presupposition, in particular following Schlenker 2008, 2009.[29] A local context is a set of worlds which represents the information locally available relative to a given syntactic environment

29 The bounded theory adopts a symmetric notion; while the role of symmetry is not crucial for present purposes, it looks to me to be well-motivated in application to indicatives. For instance, sentences with the form $(p >_i q) \wedge \neg p$ strike me as just as infelicitous as those with the reverse order, as illustrated by (21):

(21) I don't know whether Bob came to the party.

a. #But suppose that, if he didn't come, he went to work, but he did come.

b. But suppose that if he hadn't come, he would have gone to work, but he did come.

and global context: in other words, the information that could be added to that environment while being guaranteed not to change the contextual meaning of the sentence as a whole. The local context for a conditional's consequent, for instance, entails its antecedent; the local context for a right conjunct entails the left conjunct; the local context for the scope of a quantifier entails its restrictor. The bounded theory posits that epistemic modals presuppose that their accessibility relation is *local* in the sense that only local context worlds can be accessed from local context worlds. So, for instance, consider a sentence with the form ⌜If p, might not p⌝, like (20c). The locality presupposition ensures that the 'might' will only be able to access p-worlds from any local context world: in other words, its accessibility will be restricted by the antecedent. It follows that a sentence like this can never be true, as long as its locality presupposition is satisfied. Similar reasoning accounts for the infelicity of the other variants in (20).

So much for the bounded theory; further details and motivation are not important for our purposes. Instead, I want to propose a local version of the indicative constraint on a rough parallel to the bounded theory's locality presupposition.[30] In fact, we don't need to change very much. Recall that the indicative constraint says that the indicative selection function must take any context world and indicative antecedent to a context world. We need only change 'context' for 'local context'. In other words, where κ represents the local context for the conditional's antecedent, our *local indicative constraint* says that $p >_i q$ presupposes that $\forall w \in \kappa : f_i(p,w) \in \kappa$. We need to then couple this with some theory of local contexts; I will continue to assume Schlenker's theory, though other choices are compatible with our main points.

How does this answer to the motivations given above? For unembedded conditionals, the local indicative constraint is equivalent to the standard indicative con-

---

Note that adopting a symmetric theory of local contexts for the formulation of the local indicative constraint does not commit us to a symmetric theory of the projection of semantic presuppositions, which is a distinct phenomenon.

30 Boylan & Schultheis (2019) independently propose a similar but subtly different constraint, motivated on independent grounds. Their constraint captures all the data I discuss here; it is motivated by way of what they call the *Qualitative Thesis*, which our theory also validates. Interestingly, their theory doesn't validate *LI*, because selection functions in their theory are assumed to shift with local contexts. This contrast brings out the importance in the present approach of locating the indicative constraint in a presuppositional dimension, rather than letting selection functions shift intra-sententially.

straint. But things are different for embedded conditionals. Consider first sentences with the form $\neg p \wedge (p >_i q)$, which, as we have seen, are infelicitous even when embedded in such a way that p remains compatible with the global context. The local context for a right conjunct will be the global context together with the left conjunct. So, in a global context c, the local context for the conditional in $\neg p \wedge (p >_i q)$ will be $c \cap \neg p$. That means that the local indicative constraint will ensure that, for any world w' in $c \cap \neg p$, $f_i(p, w') \in (c \cap \neg p)$. But this constraint clashes with the *Success* constraint on f_i which says that $f_i(p, w')$ must be in p. So, given *Success*, there will be no way to satisfy the local indicative constraint. Crucially, this reasoning goes through whether $\neg p \wedge (p >_i q)$ is embedded or unembedded: there will be no way to satisfy the local indicative constraint in either case at any local context world. This accounts for the infelicity of sentences which embed $\neg p \wedge (p >_i q)$, like (16a). Parallel considerations will account for the infelicity of quantified sentences with the form $\forall x(p(x), \neg p(x) >_i q(x))$, as in (17a), since the local context for the conditional's antecedent here will entail the quantifier's restrictor, $\neg p(x)$. Finally, the local context for the consequent of an indicative conditional is the global context intersected with the antecedent.[31] So, in a conditional with the form $p >_i (\neg p >_i q)$ in global context c, the local context for the consequent is $c \cap p$, and so the local indicative constraint for the embedded conditional will entail that, for any world $w' \in c \cap p$, $f(\neg p, w') \in c \cap p$. This will again conflict with *Success*, accounting for the infelicity of sentences like (18a). Since subjunctive conditionals do not have a corresponding local indicative constraint, none of this reasoning will go through for subjunctives, accounting for the contrasts observed in (16)–(18). So the local indicative constraint, unlike the standard indicative constraint, accounts for the embedding data given above.

31 See Mandelkern & Romoli 2017 for this calculation in the context of Schlenker's framework. Some have proposed that subjunctive conditionals have a parallel possibility presupposition, but one which quantifies over something like circumstantially possible worlds, rather than epistemically possible ones. This is compatible with our present commitments, since circumstantial possibility, unlike epistemic possibility, is not constrained by local contexts—so even if we take on board a circumstantial possibility presupposition for subjunctives, *IE* will still not be Strawson valid for subjunctives.

Putting together Stalnaker's truth conditions with the local indicative constraint in our two-dimensional framework, we have:[32]

- $[\![p >_s q]\!]^{\kappa, f_i, f_s, w} = \langle \underline{1}, 1 \rangle$ iff $[\![q]\!]^{[\![p]\!]^{\kappa, f_i, f_s}, f_i, f_s, f_s([\![p]\!]^{\kappa, f_i, f_s}, w)} = \langle \underline{1}, 1 \rangle$

  [Stalnaker truth conditions]

- $[\![p >_i q]\!]^{\kappa, f_i, f_s, w}$

  $= \langle \underline{1}, * \rangle$ iff $\forall w' \in \kappa : f_i([\![p]\!]^{\kappa, f_i, f_s}, w') \in \kappa$

  [local indicative constraint]

  $= \langle \underline{*}, 1 \rangle$ iff $[\![q]\!]^{\kappa \cap [\![p]\!]^{\kappa, f_i, f_s}, f_i, f_s, f_i([\![p]\!]^{\kappa, f_i, f_s}, w)} = \langle \underline{1}, 1 \rangle$

  [Stalnaker truth conditions]

7 Back to *Import-Export*

The local indicative constraint was motivated on the basis of general observations about the differences between indicatives and subjunctives. Now I will return to the questions about the logic of conditionals which started us off, showing that the local indicative constraint has surprising and desirable consequences in this area: it predicts that *IE* will be Strawson valid for indicatives, not subjunctives, and that pairs that instantiate *IE* for indicatives, but not subjunctives, will be Strawson acceptance equivalent and, as a default matter, Strawson probabilistically equivalent. I take each of these points in turn, beginning with acceptance.

7.1 Acceptance equivalence

Recall that the pairs that instantiate *IE* have the form of (22b) and (22a), respectively:

(22) a. $(p \wedge q) >_i r$
b. $p >_i (q >_i r)$

Consider a context s, and suppose that the local indicative constraints of (22a) and (22b) are satisfied throughout s. Then (22a) is accepted in s (i.e. true throughout

32 As usual, I omit the world parameter to denote the set of worlds where the sentence in question is true, so $[\![p]\!]^{\kappa, f_i, f_s} = \{w : [\![p]\!]^{\kappa, f_i, f_s, w} = \langle \underline{*}, 1 \rangle\}$.

s) just in case (22b) is. I leave the proof to a footnote.[33] The key intuition is that, given the local indicative constraint, you accept either conditional just in case all the $p \wedge q$-worlds compatible with what you accept must be r-worlds. But this reasoning turns crucially on the local indicative constraint, so nothing similar follows for the subjunctive analogues of (22a) and (22b).

7.2 Strawson validity of *IE*

Very similar reasoning shows that *IE* is Strawson valid. Again, I leave the proof of this in a footnote; the intuition behind it is similar to the intuition behind the proof that the two sides of *IE* are Strawson acceptance equivalent.[34]

---

33 Suppose first that (22a) is accepted in s. Since (by *Strong Centering*) the closest $p \wedge q$-world to any $p \wedge q$-world is itself, and since (22a) is accepted in s, we know that every $p \wedge q$-world in s is an r-world. Now consider any world $w \in s$. (22b) is true at w iff $q >_i r$ is true at the closest p-world to w, call it w_p, relative to the local context $s \cap p$. By the local indicative constraint and *Success*, we know that $w_p \in s \cap p$. So the local indicative constraint on $q >_i r$ tells us that the closest q-world to w_p, call it w_q, has to be in $s \cap p$. So w_q will be a $p \wedge q$-world in s. But we know that all the $p \wedge q$-worlds in s are r-worlds. So (22b) is true at w; since w was chosen from s arbitrarily, (22b) is true throughout s. Conversely, suppose that (22b) is accepted in s. Consider any $p \wedge q$-world w in s. By *Strong Centering*, the closest p-world to w is w, and the closest q-world to w is w, so r is true at w since (22b) is true at w. That means that, given that (22b) is true at every world in s, every $p \wedge q$-world in s is an r-world. Now consider an arbitrary world $w \in s$. By the local indicative constraint, the nearest $p \wedge q$-world to w must also be in s; but then it must also be an r-world since all $p \wedge q$-worlds in s are r-worlds. So (22a) will be true at w. Since w was chosen arbitrarily from s, (22a) will be true throughout s.

34 *IE* says that the conjunction of material conditionals $((p >_i (q >_i r)) \supset ((p \wedge q) >_i r)) \wedge (((p \wedge q) >_i r) \supset (p >_i (q >_i r)))$ is always true. I prove that each material conditional is Strawson valid, which suffices to prove that the conjunction is. First consider $(p >_i (q >_i r)) \supset ((p \wedge q) >_i r)$. Suppose there is a context c, world $w \in c$, and selection functions f_i and f_s such that the presuppositions of all the indicative conditionals are satisfied at $\langle c, f_i, f_s, w \rangle$ but this material conditional is false at $\langle c, f_i, f_s, w \rangle$. Then the antecedent must be true and the consequent must be false. The local context for the consequent of a material conditional is the global context together with its antecedent. So we have $(p \wedge q) >_i r$ false at $\langle c \cap (p >_i (q >_i r)), f_i, f_s, w \rangle$. By the local indicative constraint, since $w \in c \cap (p >_i (q >_i r))$, $f_i(p \wedge q, w) \in c \cap (p >_i (q >_i r))$. But any $p \wedge q$-world that makes $p >_i (q >_i r)$ true makes r true, by *Strong Centering*; so r is true at $f_i(p \wedge q, w)$, so $(p \wedge q) >_i r$ is true after all, contrary to assumption. Next consider $((p \wedge q) >_i r) \supset (p >_i (q >_i r))$. Suppose the presuppositions of all the indicative conditionals are satisfied at $\langle c, f_i, f_s, w \rangle$ but the material conditional is false at $\langle c, f_i, f_s, w \rangle$. Then $p >_i (q >_i r)$ is false at $\langle c \cap ((p \wedge q) >_i r), f_i, f_s, w \rangle$. By the local indicative constraints, $f_i(p, w) \in c \cap ((p \wedge q) >_i r)$; and so again by the local indicative constraints, $f_i(q, f_i(p, w)) \in c \cap ((p \wedge q) >_i r) \cap p$, and so will be a $p \wedge q$-world and a $(p \wedge q) >_i r$-world and hence an r-world, so $p >_i (q >_i r)$ is true at $\langle c \cap ((p \wedge q) >_i r), f_i, f_s, w \rangle$ after all, contrary to assumption.

Note that this proof depends on treating '$\equiv$' as a conjunction of material conditionals, rather than as its own defined connective. If we treat it as its own connective, then Schlenker's local

7.3 Probability

The final, and most subtle, point I want to make concerns probability. Assuming their local indicative constraints are satisfied in a given context s, it is plausible that, as a default matter, the probability measure associated with s (i.e. which takes s as its state space) will assign the same probabilities to $(p \wedge q) >_i r$ and $p >_i (q >_i r)$.

To see this, first note that $p >_i (q >_i r)$ and $p >_i ((p \wedge q) >_i r)$ are Strawson equivalent, and thus Strawson probabilistically equivalent, on our theory.[35] Next we assume that conditionals are generally probabilistically independent of their antecedents, as well as salient *parts* of their antecedent (in roughly the sense of Yablo 2014), relative to the probability measure associated with their context. These assumptions have been extensively motivated in the literature. The first is exactly what we need to validate a limited version of the claim, from Stalnaker 1970, Adams 1975, that the probability of a conditional equals the probability of the consequent conditional on the antecedent.[36] Lewis (1976) and many following have shown that this cannot hold unrestrictedly. However, it is still plausible that it holds relative to the probability assumption associated with the same context where the conditional is

context algorithm predicts that the local context for each side is the global context, and this proof does not go through. This, however, seems unproblematic for present purposes, since, as I discuss below, the key dialectical purpose of the present point is explaining why conjunctions with the form $\Diamond((p >_i (q >_i r)) \wedge \neg((p \wedge q) >_i r))$ and vice versa cannot be entertained. Relatedly, because of the way '$\equiv$' manipulates local contexts, we do not have a deduction theorem for Strawson validity: while $(p >_i (q >_i r)) \equiv ((p \wedge q) >_i r)$ is Strawson valid, $p >_i (q >_i r)$ does not Strawson entail $(p \wedge q) >_i r$ nor vice versa. This is a potentially problematic lacuna in the theory. On the other hand, this meta-linguistic fact may be intuitively inaccessible, because of the corresponding object-language fact (the Strawson validity of *IE*), which may suffice to account for intuitions here (this would be a kind of limited error theory: the meta-language inequivalence is somehow screened off for us by the object-language equivalence).

35 Since, for world w in context c, $p >_i (q >_i r)$ is true at w iff the closest q-world w_q to the closest p-world w_p to w is in r; and w_q must be in $c \cap p$, by the local indicative constraint. $p >_i ((p \wedge q) >_i r)$ is true at w iff the closest $p \wedge q$-world $w_{p \wedge q}$ to w_p is in r. But given that w_q is the closest q-world to w_p, and is in p, and that $w_{p \wedge q}$ is the closest $p \wedge q$-world to w_p, it follows from *CSO* that $w_p = w_{p \wedge q}$. So these two conditionals are true at exactly the same context worlds, whenever their local indicative constraints are satisfied. Then we assume that Pr_s will assign the same probability to any propositions a and b provided that they have the same truth-value throughout s. Alternately, we might assume only that $Pr(a) = Pr(b)$ if a and b have both the same truth-value and the same presupposition-value at any world. The distinction between these views will matter only when p is itself a conditional.

36 As Stalnaker 1974a, Ellis 1978, Rothschild 2013 observed: then we have $Pr(p >_i q) = Pr(p >_i q | p) = \frac{Pr((p >_i q) \wedge p)}{Pr(p)} = \frac{Pr(p \wedge q)}{Pr(p)} = Pr(q|p)$.

evaluated, at least as a default matter (see Bacon 2015 for motivation and a tenability proof). The assumption can be motivated based on intuitions about the probabilities of particular cases (see Douven 2015 for extensive discussion), as well as on more theoretical grounds: the probability of an indicative conditional $p >_i q$ intuitively should depend just on what portion of the p-worlds in the context are q-worlds, not on what goes on in the $\neg p$-region. If this is correct, then parallel reasoning suggests that conditionals should also be probabilistically independent, not only of their antecedents, but also of suitable "parts" of the antecedent, i.e. suitable supersets of the antecedent, since, again, ruling out $\neg p$-worlds shouldn't affect the probability of $p >_i q$. This assumption can't hold for supersets of the antecedent in general, but plausibly it holds for salient parts of the antecedent—including, in particular, conjuncts of a conjunctive antecedent.[37] If these two default assumptions hold in a given context s, then, whenever $Pr_s(p) > 0$, $Pr_s(p >_i (q >_i r)) = Pr_s((p \wedge q) >_i r)$, assuming the conditionals' local indicative constraints are satisfied throughout s.[38] This reasoning, once again, crucially relied on the indicative constraint, so nothing similar will hold for subjunctives.

---

37 See Dorr & Hawthorne 2018, Fitelson 2019 for discussion and motivation of similar constraints. The general assumption in the neighborhood is that, where $p \subseteq x$, $Pr_s(p >_i q|x) = Pr_s(p >_i q)$. It's immediately clear that this assumption can't hold in general, but it is plausible, and possible, for it to hold for a limited set of supersets of p.

38 Proof:

i. $Pr_s(p >_i (q >_i r)) =$ by the local indicative constraint, see Footnote 35

ii. $Pr_s(p >_i ((p \wedge q) >_i r)) =$ by probabilistic independence of $p >_i ((p \wedge q) >_i r)$ and p, given $Pr_s(p) > 0$

iii. $Pr_s(p >_i ((p \wedge q) >_i r)|p) =$ by the probability calculus

iv. $\frac{Pr_s((p >_i ((p \wedge q) >_i r)) \wedge p)}{Pr_s(p)} =$ by *Strong Centering*

v. $\frac{Pr_s(p \wedge ((p \wedge q) >_i r))}{Pr_s(p)} =$ by probabilistic independence of $(p \wedge q) >_i r$ with p

vi. $\frac{Pr_s(p) * Pr_s((p \wedge q) >_i r)}{Pr_s(p)} =$

vii. $Pr_s((p \wedge q) >_i r)$.

7.4 Indistinguishability

I will now argue that, taken together, these three points predict a kind of indistinguishability for pairs which instantiate *IE*: even though *IE* is not logically valid, our theory predicts that it will be very hard to find particular cases where our intuitions about $(p \wedge q) >_i r$ differ from our intuitions about $p >_i (q >_i r)$.

First, our theory predicts that we won't be able to contrive a context where both $(p \wedge q) >_i r$ and $p >_i (q >_i r)$ can be felicitously used, but where you accept one but not the other.

One might worry that this Strawson acceptance equivalence is not enough. As long as it is possible for one of these sentences to be true and the other to be false, one might worry that we will be able to directly see that sentences of the form $(p >_i (q >_i r)) \wedge \neg((p \wedge q) >_i r)$, or of the form $\neg(p >_i (q >_i r)) \wedge ((p \wedge q) >_i r)$, will be consistent. But the second point we saw above is that, while these conjunctions are *logically* consistent, they are *Strawson* inconsistent: they cannot be both true and have their presuppositions satisfied.

One might still worry that, even if we cannot naturally talk about the possibility of $(p >_i (q >_i r)) \wedge \neg((p \wedge q) >_i r)$ or $\neg(p >_i (q >_i r)) \wedge ((p \wedge q) >_i r)$ being true, we could still assign different credences to each conditional. But this, in turn, is ruled out, assuming the defaults laid out in the last section, when these conditionals' presuppositions are satisfied. Again, those defaults cannot hold in full generality, but they will plausibly hold in enough core cases to make sense of the intuition that these seem to have the same probability.

We started out with a puzzle: we have strong indirect evidence that *IE* is invalid, for both indicative and subjunctive conditionals. But we seem to find direct counterexamples to *IE* only for subjunctives. The local indicative constraint provides a solution to this puzzle. This constraint is motivated on entirely independent grounds concerning general differences between indicatives and subjunctives. Somewhat surprisingly, this constraint predicts that the indicative pairs that instantiate the *IE*-schema will be indistinguishable in ordinary contexts where they are used. Since none of this reasoning extends to subjunctives, we account for the contrast between indicatives and subjunctives in this area.[39]

Before concluding, let me make a few big-picture points about how we avoid the collapse result above. Like *IE*-validating theories, our theory of the indicative conditional has a parameter which keeps track of subsequent indicative antecedents. But, unlike in those theories, in our theory this parameter does not provide a domain of quantification for the conditional. Instead, it provides a presuppositional *constraint* on the conditional's domain of quantification. In other words, it limits possible meanings for the embedded conditional, rather than shifting what proposition the embedded conditional expresses. This allows us to keep our logic conservative while making sense of *IE*-friendly intuitions. Crucially, we thus still validate *LI* and *Triviality*, and also avoid the collapse result above: our indicative conditional is not the material conditional. Indeed, our conditional has exactly the logic of Stalnaker's conditional, since it differs from Stalnaker's conditional only with respect to its presuppositions (its *Strawson logic* will thus be a superset of Stalnaker's logic). In particular, inferences like those from $\neg(p >_i q)$ to $p \wedge \neg q$, which are so disastrously valid on the material conditional, will be invalid for our conditional (both logically and in the Strawson sense).[40]

39 There remains a residual question of why *IE* appears to be a default assumption for subjunctives as well—after all, in many cases $p >_s (q >_s r)$ does strike us as equivalent to $(p \wedge q) >_s r$. This seems to me addressable in a broadly Stalnakerian framework: a conditional antecedent p makes salient the question of p, and thus suggests that p-ness is one of the things that the contextually salient accessibility relation will hold fixed if it can. This would be, however, a defeasible default, exactly as it appears to be in the case of subjunctives, where it appears to be overridden in the cases I gave above. A related question concerns *MP*. I will not address *MP* at length here, since the heart of my point in the first half of the paper was that *MP* is not really the key player in Gibbard's result. One could, of course, invalidate both *IE* and *MP*, which would be consistent with everything I care about in this paper; but this would be an unusual position (see Boylan & Schultheis 2019 for this approach). If we validate *MP*, as the theory I am proposing does, we must find a way to explain away McGee's counterexamples as involving some kind of context shift. The theory that I propose has some resources in this direction: insofar as we are pressured to interpret complex conditionals in an *IE*-friendly way (by the local indicative constraint on indicatives, and by general pragmatic pressures in the case of subjunctives), we may be inclined to use a different selection function to evaluate $p > (q > r)$ than we use to evaluate $q > r$, even when we evaluate them in sequence. More exploration of both these issues is needed.

40 For more exploration of the relationship between the indicative and the material conditional in a framework much like the present one, see Boylan & Schultheis 2019.

But how exactly do we block the collapse result? In particular, what does our theory say about sentences with the form $(\neg(p >_i q) \wedge q) >_i \neg(p > q)$ and close variants, which play the starring role in that result? Recall that sentences with this form are valid thanks to *LI*; whereas *IE* says that sentences with the form $(\neg(p >_i q) \wedge q) > (p >_i q)$ are instead valid. Our theory of the indicative avoids the triviality result by predicting that sentences with the first form are always true, while sentences of the second form are never (non-trivially) true. But it also makes an interesting further prediction: sentences with either of these forms cannot ever have their presuppositions satisfied. Indeed, this goes for any indicative conditional with an antecedent with the form $(\neg(p >_i q) \wedge q)$ or $(q \wedge \neg(p >_i q))$.[41] So our account predicts that these two crucial premises in our triviality result will, in the case of indicative conditionals, never have their presuppositions satisfied. And this looks right: indicatives with this form, as we noted above, sound weird. We observed this above for the indicative instantiations of our key sentences in (7a); further evidence for this generalization comes from pairs like those in (23):

(23) a. #If Bob was at the party, but it's not the case that Bob was at the party if Sue was, then…
b. If Bob had been at the party, but it's not the case that he would have been there if Sue had been, then…

The contrast between (23a) and (23b) provides another argument for our local indicative constraint. And it suggests that our theory avoids the triviality result above in a way which is not only formally coherent but also matches intuitions.

Finally: a natural question to raise at this point is why we should logically validate *LI* and only Strawson validate *IE*, rather than vice versa. After all, if our key sentences where these diverge for indicatives are both odd, then it seems like we could equally adopt an approach on which *LI* is only Strawson valid, and *IE* logically valid.[42] I want to note two responses. First, although our key conditionals

41 This is because no context world can make one of these conjunctions true and satisfy its presuppositions. Focus on the second sentence: the local context for the right conjunct will entail q; so by the indicative constraint, any context world which makes q true will be such that the closest p-world to it is a q-world, so the right conjunct will have to be false.

42 Thanks to Sam Carter for pushing this point.

in the indicative mood are indeed somewhat odd, I still think that intuition favors the predictions of *LI* over those of *IE*, as I argued in §4.2. Second, as we have seen, in the subjunctive case things look different: there our target conditionals are felicitous, and the predictions of *LI*, not *IE*, are clearly correct. So for subjunctives, it seems clear that we should logically validate *LI*, and both logically and Strawson invalidate *IE*. Insofar as we want as unified as possible a theory of indicatives and subjunctives, then, we should logically validate *LI* and logically invalidate *IE* for all conditionals, and then find a way to predict that the latter is Strawson valid for indicatives.

Having said this, let me conclude this section with an ecumenical remark. While I think the details of the account I have given here have much to recommend them, it is as much my aim to give a sketch of the general *kind* of response one might give to the puzzle I have set out, as it is to advocate the details of the present response. There are many alternatives in the neighborhood to be explored which might account for intuitions in this area in broadly similar ways.

8 Conclusion

An adequate theory of the conditional must navigate a narrow passage between, on the one hand, logical principles that seem inviolable; and, on the other, the material analysis, which is untenable. In this paper I have showed that this passage is even narrower than it seems: in particular, the only way to validate *LI*, *IE*, and *Triviality* together is with the material conditional. This result helps explain why every extant theory of the conditional which validates *IE* other than the material theory invalidates *LI*, and, indeed, invalidates its simplest and most plausible instance: *Identity*. I have argued that this violation of *LI* in general, and of *Identity* in particular, is unacceptable: after all, if p, then p! On this basis, I have argued that we must reject *IE*.

But can we find direct evidence that *IE* is invalid? In the case of subjunctive conditionals, I have argued we can: the existing literature contains striking cases which suggest that *IE* is not valid for subjunctives. But things are, surprisingly, rather different for indicative conditionals. In the case of indicatives, the same pairs of conditionals that constitute counterexamples in the subjunctive mood sound equivalent in the indicative mood.

In the second part of the paper, I developed a theory which aims to account for these subtle facts. The truth-conditions of the theory are those of Stalnaker's, and so its logic is that of Stalnaker's conditional: in particular, like Stalnaker, we invalidate *IE* and validate *LI* and *Triviality*. On top of those truth-conditions, I have proposed the local indicative constraint, which is independently motivated on the basis of a wide range of contrasts between indicatives and subjunctives. This constraint has a surprising result: even though *IE* is not logically valid for indicatives, the members of indicative pairs which instantiate *IE* will be indistinguishable in a number of ways. This lets us make sense of the lack of direct counterexamples to *IE* for indicatives, without actually predicting that it is logically valid.

There is no doubt much more to say about the differences between indicatives and subjunctives. But I am hopeful that the general strategy taken here—recruiting general differences in the presuppositions of these conditionals to account for apparent differences in their logic—will prove fruitful. My hope in particular is that this strategy will help us better understand how the conditional can occupy the narrow space between *prima facie* plausible principles like *LI* and *IE* on the one hand, and implausible theories—like the material analysis—on the other.

A Appendix A: Failures of *Identity* in McGee (1985)'s theory

On McGee (1985)'s theory, sentences are evaluated relative to three parameters. The first is a Stalnakerian *selection function* f (Stalnaker 1968, Stalnaker & Thomason 1970) from consistent propositions and worlds to worlds. The second parameter is a set of sentences Γ, which keeps track of conditional antecedents. With $\mathfrak{I}$ an atomic valuation function, we then have:

- $[\![p]\!]^{f,\Gamma,w} = 1$ if $\bigcap_{r\in\Gamma} [\![r]\!]^{f,\varnothing} = \varnothing$; else *Absurd*
- $[\![A]\!]^{f,\Gamma,w} = 1$ iff $f(\bigcap_{p\in\Gamma} [\![p]\!]^{f,\varnothing}, w) \in \mathfrak{I}(A)$ *Atom*
- $[\![\neg p]\!]^{f,\Gamma,w} = 1$ iff $[\![p]\!]^{f,\Gamma,w} = 0$ *Neg*
- $[\![p \wedge q]\!]^{f,\Gamma,w} = 1$ iff $[\![p]\!]^{f,\Gamma,w} = 1$ and $[\![q]\!]^{f,\Gamma,w} = 1$ *Conj*
- $[\![p > q]\!]^{f,\Gamma,w} = [\![q]\!]^{f,\Gamma\cup\{p\},w}$ *Cond*

Now consider an arbitrary model of McGee's semantics with at least one atomic sentence A which is true in some world and false in some other world. Choose an arbitrary world w in the model, and an arbitrary selection function f. Then:

i. $[\![(\neg(\neg A > A) \wedge A) > (\neg(\neg A > A) \wedge A)]\!]^{f,\varnothing,w} = 1$ iff

ii. $[\![(\neg(\neg A > A) \wedge A)]\!]^{f,\{(\neg(\neg A > A) \wedge A)\},w} = 1$ iff — By *Cond*

iii. $[\![\neg(\neg A > A)]\!]^{f,\{(\neg(\neg A > A) \wedge A)\},w} = 1$ and $[\![A]\!]^{f,\{(\neg(\neg A > A) \wedge A)\},w} = 1$ — By *Conj*

iv. $[\![\neg(\neg A > A)]\!]^{f,\{(\neg(\neg A > A) \wedge A)\},w} = 1$ iff

v. $[\![\neg A > A]\!]^{f,\{(\neg(\neg A > A) \wedge A)\},w} = 0$ iff — By *Neg*

vi. $[\![A]\!]^{f,\{(\neg(\neg A > A) \wedge A), \neg A\},w} = 0$ — By *Cond*

vii. $[\![A]\!]^{f,\{(\neg(\neg A > A) \wedge A), \neg A\},w} = 1$

By *Absurd*, since $[\![(\neg(\neg A > A) \wedge A)]\!]^{f,\varnothing} \cap [\![\neg A]\!]^{f,\varnothing} = \varnothing$

viii. $[\![\neg A > A]\!]^{f,\{(\neg(\neg A > A) \wedge A)\},w} = 1$ — From v–vii

ix. $[\![\neg(\neg A > A)]\!]^{f,\{(\neg(\neg A > A) \wedge A)\},w} = 0$ — From viii, v

x. $[\![(\neg(\neg A > A) \wedge A)]\!]^{f,\{(\neg(\neg A > A) \wedge A)\},w} = 0$ — From iii, ix

xi. $[\![(\neg(\neg A > A) \wedge A) > (\neg(\neg A > A) \wedge A)]\!]^{f,\varnothing,w} = 0$ — From i, x

Since w and f were chosen arbitrarily, $\ulcorner(\neg(\neg A > A) \wedge A) > (\neg(\neg A > A) \wedge A)\urcorner$ is false relative to the empty premise set together with any world and selection function in our model.

B Appendix B: *LI* + *IE* + *Triviality* lead to collapse

i. $\models (\neg(a > c) \wedge c) > \neg(a > c)$ — *LI*, since $p \wedge q \models p$

ii. $\models (c \wedge a) > c$ — *LI*, since $p \wedge q \models p$

iii. $\models c > (a > c)$ — *IE*, ii

iv. $\models \neg(a > c) > (c > (a > c))$ — *LI*, iii, since when $\models q$, then $p \models q$

v. $\models (\neg(a > c) \wedge c) > (a > c)$ — *IE*, iv

vi. $\models \neg(\neg(a > c) \wedge c)$ — *Triviality*, i, v

vii. $c \models a > c$ — vi, classical assumptions about $\wedge, \neg$

viii. $(a \wedge (a > c)) \wedge \neg c \models (a > c) \wedge (a > \neg c)$

Substituting $\neg c$ for c in vii, classical assumptions about $\wedge$

ix. $(a \wedge (a > c)) \wedge \neg c \models \neg a$ — viii, *Triviality*

x. $\models \neg((a \wedge (a > c)) \wedge \neg c)$ — ix, reductio

xi. $a \wedge (a > c) \models c$ — x, classical assumptions about $\wedge, \neg$

xi is equivalent to *MP*, so Gibbard (1981)'s result gets us the rest of the way to collapse. (Proof sketch of Gibbard's result: by *LI*, we have $\models ((a \supset c) \wedge a) > c$; by *IE*, we have $\models (a \supset c) > (a > c)$; by *MP*, we have $\models (a \supset c) \supset (a > c)$. *MP* also gives us $\models (a > c) \supset (a \supset c)$; so we have $a > c \dashv\models a \supset c$.)

References

Adams, Ernest. 1975. *The logic of conditionals*. Dordrecht.

Aloni, Maria D. 2001. *Quantification under conceptual covers*. Amsterdam: University of Amsterdam dissertation.

Arló-Costa, Horacio. 2001. Bayesian epistemology and epistemic conditionals: On the status of the Export-Import laws. *The Journal of Philosophy* 98(11). 555–593.

Arló-Costa, Horacio & Paul Egré. 2016. The logic of conditionals. In Edward N. Zalta (ed.), *The stanford encyclopedia of philosophy*, Metaphysics Research Lab, Stanford University winter 2016 edn.

Bacon, Andrew. 2015. Stalnaker's thesis in context. *Review of Symbolic Logic* 8(1). 131–163.

Bassi, Itai & Moshe E. Bar-Lev. 2018. A unified existential semantics for bare conditionals. In Robert Truswell, Chris Cummins, Caroline Heycock, Brian Rabern & Hannah Rohde (eds.), *Sinn und Bedeutung*, vol. 21, 125–142.

Bledin, Justin. 2015. *Modus Ponens* defended. *Journal of Philosophy* CXII(2). 57–83.

Bonevac, Daniel, Josh Dever & David Sosa. 2013. Conditionals and their antecedents. Manuscript, University of Texas-Austin.

Boylan, David & Ginger Schultheis. 2019. Attitudes, conditionals, and the qualitative thesis. Manuscript, NYU and Rutgers.

Ciardelli, Ivano. 2019. Indicative conditionals and graded information. Manuscript, MCMP.

Dekker, Paul. 1993. *Transsentential meditations: Ups and downs in dynamic semantics*: University of Amsterdam dissertation.

Dekker, Paul. 2008. A multi-dimensional treatment of quantification in extraordinary English. *Linguistics and Philosophy* 31(1). 101–127.

Dorr, Cian & John Hawthorne. 2013. Embedding epistemic modals. *Mind* 122(488). 867–913. http://dx.doi.org/10.1093/mind/fzt091.

Dorr, Cian & John Hawthorne. 2018. If...: A theory of conditionals. Manuscript, NYU and USC.

Dorst, Kevin. 2019. Abominable KK failures. *Mind Online First* .

Douven, Igor. 2015. *The epistemology of indicative conditionals: Formal and empirical approaches*. Cambridge: Cambridge University Press.

Edgington, Dorothy. 1986. Do conditionals have truth-conditions? *Crítica: Revista Hispanoamericana de Filosofía* 18(52). 3–39.

Ellis, Brian. 1978. A unified theory of conditionals. *Journal of Philosophical Logic* 107–124.

Etlin, David. 2008. Modus ponens revisited. Manuscript, MIT.

Field, Hartry. 2008. *Saving truth from paradox*. Oxford University Press.

Field, Hartry. 2016. Indicative conditionals, restricted quantification, and naive truth. *Review of Symbolic Logic* 9(1). 181–208. http://dx.doi.org/10.1017/S1755020315000301.

von Fintel, Kai. 1994. *Restrictions on quantifier domains*: University of Massachusetts at Amherst dissertation.

von Fintel, Kai. 1998. The presupposition of subjunctive conditionals. In Orin Percus & Uli Sauerland (eds.), *The interpretive tract*, vol. 25, 29–44. Cambridge, Massachusetts: MIT Working Papers in Linguistics. https://philpapers.org/rec/VONTPO-2.

von Fintel, Kai. 1999. NPI licensing, Strawson entailment, and context dependency. *Journal of Semantics* 16(2). 97–148.

von Fintel, Kai. 2001. Counterfactuals in a dynamic context. In *Ken Hale: A life in language*, 123–152. MIT Press. http://web.mit.edu/fintel/fintel-2001-counterfactuals.pdf.

Fitelson, Branden. 2019. How to model the epistemic probabilities of conditionals. Handout, Northeastern University.

Gibbard, Allan. 1981. Two recent theories of conditionals. In William L. Harper, Robert Stalnaker & Glenn Pearce (eds.), *Ifs: Conditionals, beliefs, decision, chance, and time*, 211–247. Dordrecht: Reidel.

Gillies, Anthony. 2004. Epistemic conditionals and conditional epistemics. *Noûs* 38(4). 585–616. http://dx.doi.org/10.1111/j.0029-4624.2004.00485.x.

Gillies, Anthony. 2007. Counterfactual scorekeeping. *Linguistics and Philosophy* 30(3). 329–360. http://dx.doi.org/10.1007/s10988-007-9018-6.

Gillies, Anthony. 2009. On truth conditions for *If* (but not quite only *If*). *Philosophical Review* 118(3). 325–349.

Grice, Paul. 1989. *Studies in the way of words*. Cambridge: Harvard University Press.

Groenendijk, Jeroen, Martin Stokhof & Frank Veltman. 1996. Coreference and modality. In Shalom Lappin (ed.), *Handbook of contemporary semantic theory*, 179–216. Oxford: Blackwell.

Harper, W. L. 1976. Ramsey test conditionals and iterated belief change (A response to Stalnaker). In W. L. Harper & C. A. Hooker (eds.), *Foundations of probability theory, statistical inference, and statistical theories of science*, vol. 6a The University of Western Ontario Series in Philosophy of Science, 117–135. Springer, Dordrecht. http://dx.doi.org/10.1007/978-94-010-1853-1_6.

Hawthorne, John. 2005. Chance and counterfactuals. *Philosophy and Phenomenological Research* 70. 396–405.

Herburger, Elena. 2018. Bare conditionals in the red. *Linguistics and Philosophy* http://dx.doi.org/10.1007/s10988-018-9242-2.

Herzberger, Hans G. 1973. Dimensions of truth. *Journal of Philosophical Logic* 2(4). 535–556.

Holguín, Ben. 2018. Knowledge in the face of conspiracy conditionals. Manuscript, NYU.

Jackson, Frank. 1987. *Conditionals*. Oxford: Blackwell.

Jenny, Matthias. 2016. Counterpossibles in science: The case of relative computability. *Noûs* 52(3). 530–560. http://dx.doi.org/10.1111/nous.12177.

Karttunen, Lauri & Stanley Peters. 1979. Conventional implicatures in Montague grammar. In *Syntax and semantics 11: Presupposition*, New York: Academic Press.

Khoo, Justin. 2013. A note on Gibbard's proof. *Philosophical Studies* 166(1). 153–164.

Khoo, Justin & Matthew Mandelkern. 2018. Triviality results and the relationship between logical and natural languages. *Mind* 128(510). 485–526. http://dx.doi.org/10.1093/mind/fzy006.

Klinedinst, Nathan & Daniel Rothschild. 2012. Connectives without truth tables. *Natural Language Semantics* 20(2). 137–175.

Kolodny, Niko & John MacFarlane. 2010. Ifs and oughts. *Journal of Philosophy* 107. 115–143.

Kratzer, Angelika. 1981. The notional category of modality. In H. Eikmeyer & H. Rieser (eds.), *Words, worlds, and contexts: New approaches in word semantics*, 38–74. de Gruyter.

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

Kratzer, Angelika. 1991. Modality. In Arnim von Stechow & Dieter Wunderlich (eds.), *Semantics: An international handbook of contemporary research*, 639–650. Berlin: de Gruyter.

Lewis, David. 1973. *Counterfactuals*. Oxford: Blackwell.

Lewis, David. 1976. Probabilities of conditionals and conditional probabilities. *The Philosophical Review* 85(3). 297–315. http://dx.doi.org/10.2307/2184045.

MacFarlane, John. 2011. Epistemic modals are assessment sensitive. In Andy Egan & Brian Weatherson (eds.), *Epistemic modality*, 144–177. Oxford University Press.

Mandelkern, Matthew. 2018. Import-export and 'and'. *Philosophy and Phenomenological Research* http://dx.doi.org/10.1111/phpr.12513.

Mandelkern, Matthew. 2019. Bounded modality. *The Philosophical Review* 128(1). 1–61. http://dx.doi.org/10.1215/00318108-7213001.

Mandelkern, Matthew & Jacopo Romoli. 2017. Parsing and presuppositions in the calculation of local contexts. *Semantics and Pragmatics* 10(17). http://dx.doi.org/10.3765/sp.10.7.

McGee, Vann. 1985. A counterexample to modus ponens. *The Journal of Philosophy* 82(9). 462–471.

McGee, Vann. 1989. Conditional probabilities and compounds of conditionals. *The Philosophical Review* 98(4). 485–541.

Ninan, Dilip. 2018. Quantification and epistemic modality. *The Philosophical Review* 127(2). 433–485. http://dx.doi.org/10.1215/00318108-6973010.

Ramsey, Frank P. 1978/1931. General propositions and causality. In D.H. Mellor (ed.), *Foundations: Essays in philosophy, logic, mathematics and economics*, 133–51. Routledge and Kegan Paul.

Rothschild, Daniel. 2013. Do indicative conditionals express propositions? *Noûs* 47(1). 49–68.

Schlenker, Philippe. 2008. Be articulate: a pragmatic theory of presupposition projection. *Theoretical Linguistics* 34(3). 157–212. http://dx.doi.org/10.1515/THLI.2008.013.

Schlenker, Philippe. 2009. Local contexts. *Semantics and Pragmatics* 2(3). 1–78. http://dx.doi.org/10.3765/sp.2.3.

Sextus Empiricus. 2005. *Against the logicians;* Richard Bett, ed. Cambridge University Press.

Silk, Alex. 2017. How to embed an epistemic modal: Attitude problems and other defects of character. *Philosophical Studies* 174. 1773–1799.

Soames, Scott. 1989. Presupposition. In Dov Gabbay & Franz Guenthner (eds.), *Handbook of philosophical logic, volume iv: Topics in the philosophy of language*, 553–616. Kluwer Academic Publishers.

Stalnaker, Robert. 1968. A theory of conditionals. In Nicholas Rescher (ed.), *Studies in logical theory*, 98–112. Oxford: Blackwell. https://philpapers.org/rec/STAATO-5.

Stalnaker, Robert. 1974a. Letter to Bas Van Fraassen.

Stalnaker, Robert. 1974b. Pragmatic presuppositions. In Milton K. Munitz & Peter Unger (eds.), *Semantics and philosophy*, 197–213. New York: New York University Press. http://dx.doi.org/10.1093/0198237073.003.0003.

Stalnaker, Robert. 1975. Indicative conditionals. *Philosophia* 5(3). 269–86. https://philpapers.org/rec/STAIC.

Stalnaker, Robert. 1978. Assertion. In Peter Cole (ed.), *Syntax and semantics*, vol. 9, 315–322. New York: Academic Press.

Stalnaker, Robert C. 1970. Probability and conditionals. *Philosophy of Science* 37(1). 64–80.

Stalnaker, Robert C. & Richmond H. Thomason. 1970. A semantic analysis of conditional logic. *Theoria* 36(1). 23–42. http://dx.doi.org/10.1111/j.1755-2567.1970.tb00408.x.

Starr, William. 2014. A uniform theory of conditionals. *Journal of Philosophical Logic* 33(6). 1019–1064.

Stojnić, Una. 2016. One's *Modus Ponens*: Modality, coherence, and logic. *Philosophy and Phenomenological Research* 93(3). 1–48.

Strawson, Peter. 1952. *Introduction to logical theory*. London: Methuen.

Sudo, Yasutada. 2012. *On the semantics of phi features on pronouns*: Massachusetts Institute of Technology dissertation.

Veltman, Frank. 1996. Defaults in update semantics. *Journal of Philosophical Logic* 25(3). 221–261.

Warmbrod, Ken. 1981a. Counterfactuals and substitution of equivalent antecedents. *Journal of Philosophical Logic* 10. 267–289.

Warmbrod, Ken. 1981b. An indexical theory of conditionals. *Dialogue: Canadian Philosophical Review* 20(644-664).

van Wijnbergen-Huitink, Janneke, Shira Elqayam & David E. Over. 2014. The probability of iterated conditionals. *Cognitive Science* 1–16. http://dx.doi.org/10.1111/cogs.12169.

Williamson, Timothy. 2019. Suppose and tell: The semantics and heuristics of conditionals. Manuscript, New College, Oxford.

Yablo, Stephen. 2014. *Aboutness*. Princeton University Press.

Yalcin, Seth. 2007. Epistemic modals. *Mind* 116(464). 983–1026. http://dx.doi.org/10.1093/mind/fzm983.