



THE JOURNAL OF PHILOSOPHY

VOLUME CXVIII, NO. 12, DECEMBER 2021

IF P, THEN P!*

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*¹

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 great deal of controversy about what kind of following is involved. But plausibly, on any reasonable notion of "following," any sentence follows 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" will be logical truths.

In the first part of this paper, I show that, despite the overwhelming plausibility of this *Identity* principle, a wide variety of theories of

*I am grateful for a great deal of very helpful feedback on this paper from audiences at Oxford, the Munich Center for Mathematical Philosophy, NYU, UCL, and USC; audiences and reviewers for the 22nd Amsterdam Colloquium; and from Andrew Bacon, Kyle Blumberg, Susanne Bobzien, David Boylan, Fabrizio Cariani, Sam Carter, Ivano Ciardelli, Nilanjan Das, Kevin Dorst, Paul Egré, Kit Fine, Branden Fitelson, Vera Flocke, Simon Goldstein, Jeremy Goodman, Irene Heim, Benj Hellie, Ben Holguín, Justin Khoo, Cameron Domenico Kirk-Giannini, Angelika Kratzer, Vann McGee, Eliot Michaelson, Kurt Norlin, Jonathan Phillips, Milo Phillips-Brown, Richard Roth, Ian Rumfitt, Bernhard Salow, Paolo Santorio, Philippe Schlenker, Ginger Schultheis, Robert Stalnaker, and Frank Veltman. Special thanks to Daniel Rothschild for extensive discussion; to Cian Dorr, who had an essential influence on the ideas in the second half of the paper; and to a referee for this JOURNAL for very helpful comments that led to many improvements.

¹Sextus Empiricus, *Against the Logicians*, ed. Richard Bett (Cambridge, UK: Cambridge University Press, 2005).

the conditional invalidate it. I then argue that the culprit behind this failure is the *Import-Export* principle, which says that \lceil If p , then if q , then r \rceil and \lceil If p and q , then r \rceil are invariably equivalent. I show that there is a deep and surprising tension between *Import-Export*, on the one hand, and *Identity*, on the other: given two very weak, nearly universally accepted background principles, the only way to validate both *Import-Export* and *Identity* is with the material conditional. In light of the overwhelming plausibility of *Identity*, and the implausibility of equating ‘if. . . then’ with the material conditional, I argue 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 in the conditional’s local context. 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.

I. A CRISIS OF IDENTITY

Identity, again, says that sentences with the form \lceil If p , then p \rceil are logical truths. *Identity* is one of the most natural, and least controversial, principles in the logic of the conditional. Arló-Costa and Egré² call it “constitutive of the very notion of conditional.” This seems correct: to argue for it, one can’t do much better than repeat the title of this paper: if p , then p ! 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 all those (apart from the material conditional) which validate the *Import-Export* principle.

I will work with a standard propositional language with atoms A, B, C, \dots , connectives ‘ \wedge ’ and ‘ \vee ’; negation ‘ \neg ’; the material conditional ‘ \supset ’ ($p \supset q$ is equivalent to $\neg p \vee q$), and the material biconditional ‘ \equiv ’ ($p \equiv q$ is equivalent to $(p \supset q) \wedge (q \supset p)$). Finally, we have a conditional connective ‘ $>$ ’; later in the paper I will distinguish

²Horacio Arló-Costa and Paul Egré, “The Logic of Conditionals,” in Edward N. Zalta, ed., *The Stanford Encyclopedia of Philosophy* (2016).

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. Where Γ is a set of sentences of our language, ' $\Gamma \models p$ ' means that Γ semantically entails p , in the standard classical sense that p is true in every world in every intended model where all the elements of Γ are true. I will assume this classical notion of entailment throughout. The use of a formal language is just to facilitate discussion, so ' $p > q$ ' is just an abbreviation of ' \ulcorner If p , then $q\urcorner$ '.³

The two principles which will play a central role in what follows are, again:

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

Identity is self-explanatory. *IE* is a bit more complicated. It says, in essence, 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, then we will win.
- b. If the coin is flipped and it lands heads, then we will win.

And likewise for the subjunctive version:

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

Both of these principles are *prima facie* plausible. The plausibility of *Identity* is, I take it, manifest; the plausibility of *IE* comes, *inter alia*, 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.⁴

³Angelika Kratzer, "Conditionals," *Proceedings from the Annual Meeting of the Chicago Linguistic Society*, xxii, 2 (1986): 1–15, famously argues that it is a mistake to treat 'if' as a two-place connective, as I do. But Justin Khoo, "A Note on Gibbard's Proof," *Philosophical Studies*, clxvi, 1 (December 2013): 153–64, 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.

⁴*IE* is the conjunction of two principles, *Importation*: $\models (p > (q > r)) \supset ((p \wedge q) > r)$; and *Exportation*: $\models ((p \wedge q) > r) \supset (p > (q > r))$. Both directions play a role

With this background on the table, let me turn to the central claim of this section: that a wide range of existing theories of the conditional which validate *IE* also invalidate *Identity*. Indeed, this is true of all the theories I know of, apart from the material analysis.

To understand why this is, 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,⁵ 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 section v 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—the closest q -world to the closest p -world need not be the same as the closest $p \wedge q$ -world—and so *IE* is invalid on this theory. 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’s framework,⁶ conditional antecedents are added sequentially to a set, and the consequent is then evaluated at the closest world where all the sentences in that set are true (see the first appendix for a more careful exposition). In the restrictor framework,⁷ conditional antecedents are similarly added to the value of a modal base function which takes each

in the proof below. A similar proof, which I give in Matthew Mandelkern, “Crises of *Identity*,” in Julian J. Schlöder, Dean McHugh, and Floris Roelofsen, eds., *Proceedings of the 22nd Amsterdam Colloquium* (Amsterdam: University of Amsterdam, 2019), pp. 279–88 (a proceedings paper which this paper extends), only relies on *Exportation* (but is more committal in some other ways). That version of the proof suggests that we could live with *Importation* but not *Exportation*; see Kurt Norlin, “Acceptance, Certainty, and Indicative Conditionals,” unpublished manuscript (2020), for a system that validates *Identity*, *Importation* in full generality, but not *Exportation* when the first antecedent is a conditional. Both directions, however, appear to fail for subjunctive conditionals, as I discuss, which suggests to me that we do not want to semantically validate either.

⁵Robert Stalnaker, “A Theory of Conditionals,” in Nicholas Rescher, ed., *Studies in Logical Theory* (Oxford: Blackwell, 1968), pp. 98–112.

⁶Vann McGee, “A Counterexample to Modus Ponens,” this JOURNAL, LXXXII, 9 (September 1985): 462–71.

⁷Angelika Kratzer, “The Notional Category of Modality,” in Hans-Jürgen Eikmeyer and Hannes Rieser, eds., *Words, Worlds, and Contexts: New Approaches to Word Semantics* (Berlin: W. de Gruyter, 1981); Angelika Kratzer, “Modality,” in Arnim von Stechow and

world to a set of propositions, which in turn provides the domain of quantification for evaluating the consequent. There are still other approaches, in dynamic frameworks⁸ and strict frameworks.⁹ These differ greatly in detail, but all these theories have some parameter which is in the business of 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.¹⁰

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 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 section III.1, we get even more concrete, looking at conditionals in natural language with this form).¹¹ This has the form $p > p$. Now consider

Dieter Wunderlich, eds., *Semantics: An International Handbook of Contemporary Research* (Berlin: W. de Gruyter, 1991), pp. 639–50; and Kai von Fintel, *Restrictions on Quantifier Domains*, PhD diss. (University of Massachusetts at Amherst, 1994).

⁸Paul Dekker, *Transsentential Meditations: Ups and Downs in Dynamic Semantics*, PhD diss. (University of Amsterdam, 1993); Anthony Gillies, “Epistemic Conditionals and Conditional Epistemics,” *Noûs*, xxxviii, 4 (December 2004): 585–616; and William Starr, “A Uniform Theory of Conditionals,” *Journal of Philosophical Logic*, xxxiii, 6 (December 2014): 1019–64.

⁹Anthony Gillies, “On Truth Conditions for *If* (But Not Quite Only *If*),” *The Philosophical Review*, cxviii, 3 (July 2009): 325–49.

¹⁰As Justin Khoo and Matthew Mandelkern, “Triviality Results and the Relationship between Logical and Natural Languages,” *Mind*, cxxviii, 510 (April 2019): 485–526; and Matthew Mandelkern, “Import-Export and ‘And’,” *Philosophy and Phenomenological Research*, c, 1 (January 2020): 118–35, 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.

¹¹If we restrict our language so that conditionals only have non-conditional antecedents, then we can unproblematically validate *IE* and *Identity* together (as in Vann

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 in the consequent. The consequent, again, is $(\neg(A > B) \wedge B)$, and hence entails that the conditional $A > B$ is false. The problem is that if the domain of worlds which matter for evaluating the conditional includes only *B*-worlds, then this conditional, on any reasonable theory of the conditional, cannot 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 can easily be).

This gives a sense of why theories which validate *IE* generally invalidate *Identity*. In the appendix, I go through this reasoning in more detail in the context of McGee's theory.¹² For now, the crucial point is that existing theories of the conditional that validate *IE* (apart from the material conditional) dramatically invalidate *Identity*. What is more, that means that the internal negation of these sentences are logically true: some sentences with the form $p > \neg p$ are logical truths according to these theories, even when *p* is possible.

McGee, "Conditional Probabilities and Compounds of Conditionals," *The Philosophical Review*, xcvi, 4 (October 1989): 485–541; Horacio Arló-Costa, "Bayesian Epistemology and Epistemic Conditionals: On the Status of the Export-Import Laws," this JOURNAL, xcvi, 11 (November 2001): 555–93; and Ivano Ciardelli, "Indicative Conditionals and Graded Information," *Journal of Philosophical Logic*, xlix, 3 (June 2020): 509–49. But this restriction is unmotivated: there is no problem embedding either indicative or subjunctive conditionals in conditional antecedents, and we can, and should, ask questions about their logic (see Bas van Fraassen, "Probabilities of Conditionals," in W. L. Harper and C. A. Hooker, eds., *Foundations of Probability Theory, Statistical Inference, and Statistical Theories of Science*, vol. 1 (Dordrecht, the Netherlands: D. Reidel, 1976), pp. 261–308, at p. 301; and Andrew Bacon, "Stalnaker's Thesis in Context," *Review of Symbolic Logic*, viii, 1 (March 2015): 131–63, for similar conclusions).

¹²Von Fintel'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 (see von Fintel, "Restrictions on Quantifier Domains," *op. cit.*). Increased flexibility does not, however, much improve the situation for these theories: *Identity* still fails for precisely the same reason as it does in theories which validate *IE* in general (though the situation is slightly improved in that the relevant sentences with the surface form $p > \neg p$ will not be valid, either, insofar as they have a non-coindexed reading).

II. THE CULPRIT

Is it an accident that existing theories of the conditional which validate *IE* invalidate *Identity*? Or 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. On the contrary, there is a deep tension between *IE* and *Identity*. The discussion in the last section already pointed toward this tension; in this section I will develop that discussion more precisely, showing that, provided we take on board two weak background assumptions that seem beyond serious doubt, the material conditional is the only conditional which validates both *IE* and *Identity*.

Indeed, the material conditional is the only conditional I know of which validates both *IE* and *Identity* (in a broadly classical setting).¹³ But there is overwhelming evidence that the natural language conditional ‘If... then...’ is not the material conditional. A quick way to see the implausibility of the material analysis is that, since $p \supset q$ is equivalent to $\neg p \vee q$, $\neg(p \supset q)$ is equivalent to the conjunction $p \wedge \neg q$. But it is clear that the negation of the natural language conditional $p > q$ is not equivalent to $p \wedge \neg q$. For instance, ‘It’s not the case that, if Patch had been a rabbit, she would have been a rodent’ and ‘It’s

¹³ De Finettian trivalent theories validate both principles under certain non-classical notions of logical entailment, though not in the classical sense of preservation of truth (see Bruno de Finetti, “La Logique de la Probabilité,” in *Actes du Congrès International de Philosophie Scientifique*, vol. 4 (Paris: Hermann Editeurs, 1936), pp. 1–9; Paul Egré, Lorenzo Rossi, and Jan Sprenger, “De Finettian Logics of Indicative Conditionals Part I: Trivalent Semantics and Validity,” *Journal of Philosophical Logic*, 1, 2 (April 2021): 187–213; Paul Egré, Lorenzo Rossi, and Jan Sprenger, “De Finettian Logics of Indicative Conditionals Part II: Proof Theory and Algebraic Semantics,” *Journal of Philosophical Logic*, 1, 2 (April 2021): 215–47; and Paul Egré, Lorenzo Rossi, and Jan Sprenger, “Gibbardian Collapse and Trivalent Conditionals,” forthcoming in S. Kaufmann, D. Over, and G. Sharma, eds., *Conditionals: Logic, Linguistics, and Psychology* (New York: Palgrave Macmillan, 2021), for extensive recent discussion). Those theories avoid the collapse result below by invalidating *Ad Falsum*. While there are important differences between the trivalent approach and the one I propose below, from one perspective there is a striking similarity: both depend on countenancing semantic distinctions beyond true and false. A central advantage of my approach over the trivalent truth-functional one is that it extends more naturally to a unified account of indicatives and subjunctives, one which captures failures of *IE* for subjunctives. (Thanks to Paul Egré, Lorenzo Rossi, and Jan Sprenger for helpful discussion on this point.)

Another non-classical approach worth mentioning is the intuitionistic one. The standard intuitionistic conditional validates *Identity*, *Mon*, *IE*, and *Ad Falsum*, but only one direction of collapse: where ‘ \rightsquigarrow ’ is the intuitionistic conditional, we have $\neg p \vee q \models p \rightsquigarrow q$ but $p \rightsquigarrow q \not\models \neg p \vee q$. In a classical setting, $\neg p \vee q \models p \rightsquigarrow q$ still yields the problematic result that $\neg(p \rightsquigarrow q) \models p \wedge \neg q$, but this does not hold in the intuitionistic setting. Still, in the intuitionistic setting we have $\neg(p \rightsquigarrow q) \models \neg\neg p \wedge \neg q$, so it is not clear how much of an improvement this is over the material conditional. Thanks to Ivano Ciardelli and Frank Veltman for helpful discussion on this point.

not the case that, if Patch is a rabbit, she is a rodent' are both clearly true, thanks just to taxonomic facts, regardless of whether Patch is a rabbit. So neither of these conditionals is equivalent to 'Patch is a rabbit and not a rodent', *pace* the material view.¹⁴ That it is possible to validate both *Identity* 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 *Identity* and *IE*, it is not the only logically possible one. However, I will argue that there is no *plausible* way to validate *Identity* and *IE* together. For, provided we take on two very weak, almost universally accepted background principles, the material conditional is indeed the only connective which validates both *Identity* and *IE*.

The first principle is a very weak monotonicity principle, which says that, if $p > p$ is a logical truth, then if p logically entails q , then $p > q$ is a logical truth as well:

- *Very Weak Monotonicity (Mon)*: if $\models (p > p)$ and $p \models q$, then $\models (p > q)$

Mon is a very weak corollary of the much more general principle that conditionals are monotone in their consequents: that is, that if q entails r , then $p > q$ entails $p > r$. This principle is very plausible: if q entails r , then r is true whenever q is; and so if q holds if p does, then surely r holds if p does. This more general principle is validated by every theory of the conditional I know of; and thus *Mon* is as well.¹⁵

Our final principle says that, if $p > q$ and $p > \neg q$ are both true, then p is false. I will call the principle *Ad Falsum*:

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

I will return in a moment to the motivation for *Ad Falsum*. For now, let me briefly summarize the reasoning which shows that the only connective which validates *Identity*, *Mon*, *IE*, and *Ad Falsum* is the material conditional; this reasoning is spelled out in more detail in the second appendix. *Identity* and *Mon* together entail that any conditional with the form of (3) is logically true:

¹⁴One 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 $p \wedge \neg q$ from $\neg(p > q)$: Gricean tools are apt for explaining how inferences get amplified, but are not generally useful for explaining how logical entailments get blocked.

¹⁵*Mon* and *Identity* are together equivalent to the more familiar *Logical Implication (LI)* principle, which says that, whenever p entails q , $p > q$ is a logical truth. It is helpful to keep these ingredient principles separate, however, since our target theories invalidate *Identity* but not *Mon*.

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

That is because (3) has the form $(a \wedge b) > a$, so the antecedent logically entails the consequent. *IE*, *Identity*, and *Mon* together entail that the internal negation of this conditional, in (4), is logically true:

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

That is because (4) is equivalent, by *IE*, to $\neg(p > q) > (q > (p > q))$, whose consequent $q > (p > q)$ is, by *IE*, equivalent to $(q \wedge p) > q$, and thus a logical truth; and by *Identity* and *Mon*, any conditional whose consequent is a logical truth is itself a logical truth. Thus it follows from *Ad Falsum* that the antecedent of (3) and (4)—namely, $(\neg(p > q) \wedge q)$ —is logically false. But, given *Ad Falsum*, if $(\neg(p > q) \wedge q)$ is logically false, then we can show that $p > q$ entails $p \supset q$.

Broadly similar reasoning lets us show that $p \supset q$ entails $p > q$. We derive $((p \supset q) \wedge \neg(p > q)) > (p > q)$ from *Identity* and *Mon* via *IE*; and we get $((p \supset q) \wedge \neg(p > q)) > \neg(p > q)$ from *Identity* and *Mon*. But then by *Ad Falsum*, $((p \supset q) \wedge \neg(p > q))$ is logically false, which means that $p \supset q \models p > q$.

And so our conditional collapses to the material conditional: $p > q \models p \supset q$ (again, see the second appendix for details).

In sum: any conditional which validates *IE*, *Identity*, *Mon*, and *Ad Falsum* 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...’.

II.1. Relation to Existing Results. Before discussing how to respond to this result, let me briefly situate it in relation to a set of famous results by Dale and Gibbard,¹⁶ which showed that *Identity*, *Mon*, *IE*, and *Modus Ponens* (*MP*, which says $\{p, p > q\} \models q$) can be jointly validated only by the material conditional.¹⁷ I have shown that the same result follows if we replace *MP* with *Ad Falsum*. And *Ad Falsum* is, by itself, strictly weaker than *MP*: any theory of the conditional which

¹⁶A. J. Dale, “A Defence of Material Implication,” *Analysis*, xxxiv, 3 (January 1974): 91–95; Tony Dale, “A Natural Deduction System for ‘If Then’,” *Logique et Analyse*, xxii, 87 (September 1979): 339–45; and Allan Gibbard, “Two Recent Theories of Conditionals,” in W. L. Harper, R. Stalnaker, and G. Pearce, eds., *Ifs: Conditionals, Beliefs, Decision, Chance, and Time* (Dordrecht, the Netherlands: D. Reidel, 1981), pp. 211–47.

¹⁷The result is usually attributed to Gibbard alone, but, as Kurt Norlin has helpfully pointed out to me, Dale’s papers make essentially the same point and were published some years before, albeit with a different conclusion. See also Peter Gibbins, “Material Implication: A Variant of the Dale Defence,” *Logique et Analyse*, xxii, 88 (December 1979): 447–52.

validates *MP* validates *Ad Falsum*, but not vice versa.¹⁸ So the present result strengthens the Dale/Gibbard result by replacing *MP* with a much weaker principle. I will presently argue, moreover, that *Ad Falsum* has a much firmer dialectical status than *MP*, since, while there are intuitive counterexamples to *MP*, there are not, as far as I know, to *Ad Falsum*. If that is right, then this strengthening of Dale/Gibbard's result shows that there is a tension that has been missed in the response to that result: there is a fundamental tension, not just between *IE* and *MP* (as that result showed), but also between *IE* and *Identity*—which, in turn, suggests that things are much worse for *IE* than they have appeared.¹⁹

III. RESPONSES

We have identified a tension between *IE*, *Ad Falsum*, *Mon*, and *Identity*. We cannot validate all of them; so which one should we reject?²⁰

III.1. Ad Falsum. Consider first *Ad Falsum*. The most direct evidence for *Ad Falsum* 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*, then *q*, and if *p*, then not *q*; from which we can conclude that *p* is false. This reasoning, however, is only valid if *Ad Falsum* is. While this reasoning is most at home in mathematical and logical contexts, it also seems perfectly valid in non-mathematical contexts, as in Gibbard's famous Sly Pete case.²¹ In that case, we learn both 'If Pete called, he won' and 'If Pete called, he lost' and can conclude with perfect confidence that he did not call.

¹⁸ *MP* entails *Ad Falsum*. Assume $(p > q) \wedge (p > \neg q)$ for conditional proof. Assume *p* for reductio. By *MP*, we can infer both *q* and $\neg q$. By reductio, we conclude $\neg p$. Discharging conditional proof, we have $\models ((p > q) \wedge (p > \neg q)) \supset \neg p$, and hence by classical assumptions, $\{p > q, p > \neg q\} \models \neg p$. By contrast, *Ad Falsum* does not entail *MP*, as witnessed by the existence of systems like McGee's which validate the former but not the latter.

¹⁹ Branden Fitelson, "A New Gibbardian Collapse Theorem for the Indicative Conditional," unpublished manuscript (Northeastern University, 2020), gives a different, very interesting strengthening of the Dale/Gibbard result. The key difference to mine is that Fitelson, following Dale/Gibbard, encodes the assumption that the indicative $p > q$ entails the "logical" conditional $p \rightarrow q$. However, Fitelson does not assume the latter is material, and in particular does not assume that it satisfies *MP*, and in that respect is similar to our result.

²⁰ All these principles might fail once we admit semantic vocabulary ('true', 'false', and so on) into our language, due to semantic paradoxes. Having said that, I think that there is much to be gained by initially developing a theory for a fragment free of semantic vocabulary, in the hopes that a theory that incorporates semantic vocabulary will be able to build directly on it (see Hartry Field, "Indicative Conditionals, Restricted Quantification, and Naive Truth," *Review of Symbolic Logic*, ix, 1 (March 2016): 181–208, for this kind of approach). And of course, these schemata are only valid if language like pronouns, implicit temporal and locative indexing, and so on, remains fixed.

²¹ Gibbard, "Two Recent Theories of Conditionals," *op. cit.*

Another way to motivate *Ad Falsum* is by way of two more general principles which are almost universally accepted, and which entail *Ad Falsum*. The first is the *Agglomeration* principle, which says that $p > q$ and $p > r$ together entail $p > (q \wedge r)$. *Agglomeration* is, in Hawthorne's words, "overwhelmingly intuitive":²² 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 (for example, we can infer 'If it rains, the picnic will be canceled and the parade will be canceled' from 'If it rains, the picnic will be canceled' and 'If it rains, the parade will be canceled').²³ The second principle says that, if $p > \perp$ is true, then p is false, where \perp is any contradiction. This principle is, again, very compelling: if $p > \perp$ is true, then \perp must in some sense follow from p ; but \perp 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 obviously entail *Ad Falsum*.

Finally, *Ad Falsum* 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 *Ad Falsum*.

Thus it is very hard to reject *Ad Falsum*. Again, nearly every theory of the conditional validates it, even those theories which reject *MP*.²⁴ And indeed, while there has been a serious case made against *MP* by McGee,²⁵ there has been no serious case made against *Ad Falsum*; and

²² John Hawthorne, "Chance and Counterfactuals," *Philosophy and Phenomenological Research*, LXX, 2 (March 2005): 396–405.

²³ The only theories I know of which invalidate *Agglomeration* treat the conditional as an *existential* operator, so that $p > q$ says *some* accessible p -world is a q -world (Itai Bassi and Moshe E. Bar-Lev, "A Unified Existential Semantics for Bare Conditionals," in Robert Truswell et al., eds., *Proceedings of Sinn und Bedeutung 21: Volume 1* (2018), pp. 125–42; and Elena Herburger, "Bare Conditionals in the Red," *Linguistics and Philosophy*, XLII, 2 (April 2019): 131–75). This view does not, however, strike me as very plausible.

²⁴ The only theory I know of, apart from existential ones, that invalidates *Ad Falsum* is the theory that comprises the conditional from Niko Kolodny and John MacFarlane, "Ifs and Oughts," this JOURNAL, CVII, 3 (March 2010): 115–43, together with the semantics for epistemic modals from Seth Yalcin, "Epistemic Modals," *Mind*, cxvi, 464 (November 2007): 983–1026; and John MacFarlane, "Epistemic Modals Are Assessment Sensitive," in Andy Egan and Brian Weatherson, eds., *Epistemic Modality* (Oxford: Oxford University Press, 2011), pp. 144–77, as Ivano Ciardelli has pointed out to me. In particular, on that theory $(\neg(p > q) \wedge q) > r$ and $(\neg(p > q) \wedge q) > \neg r$ are logical truths for any p, q, r , but $\neg(\neg(p > q) \wedge q)$ is not a logical truth. This is intriguing, but at least somewhat orthogonal to present issues, since that theory also invalidates *IE*.

²⁵ McGee, "A Counterexample to Modus Ponens," *op. cit.*

I do not see any way to convert the standard case against *MP*, based on complex conditional consequents, into a case against *Ad Falsum*.

III.2. Very Weak Monotonicity. So let us assume that *Ad Falsum* is not the culprit, and explore instead the possibility of rejecting *Mon*, *Identity*, or *IE*.

I do not know of any theories that reject *Mon*. *Mon* is, again, a very weak corollary of a much stronger principle, which says that, whenever q logically entails r , $p > q$ logically entails $p > r$. This principle can be motivated on the basis of inferences like this:

- (5) If Sue goes the picnic, then Mark will be sad and Liz will be happy.
 \hookrightarrow If Sue goes to the picnic, then Mark will be sad.
- (6) If Sue had gone to the picnic, then Mark would have been sad and Liz would have been happy.
 \hookrightarrow If Sue had gone to the picnic, then Mark would have been sad.

It is not just that *Mon* has never been explicitly questioned in the literature; it has not even been *implicitly* questioned, insofar as it is valid on every theory of the conditional that has been put forward (to my knowledge), including those which invalidate *Identity* and *MP*. This puts *Mon* in a slightly different category from *Identity*: even though *Identity* is, to my mind, even more obviously valid than *Mon*, *Identity* has been invalidated by a range of theories, as we have seen.

I thus do not think there is a case to be made for rejecting *Mon*.

III.3. *Identity*. This leaves *Identity* and *IE*. Unlike rejecting *Ad Falsum* or *Mon*, rejecting *Identity* and *IE* have been taken seriously in the literature, the former implicitly, the latter explicitly.

Start with *Identity*. While, as we have seen, a broad range of theories in fact do invalidate *Identity*, there are very few explicit arguments against it.²⁶

²⁶ One exception is Sextus Empiricus's famously obscure "emphasis" account, which invalidates *Identity* because "it is impossible that anything be included in itself" (*PH* 2.112; see Yale Weiss, "Sextus Empiricus' Fourth Conditional and Containment Logic," *History and Philosophy of Logic*, XL, 4 (2019): 307–22, for a semantic reconstruction in a truth-maker framework). Certain relevance logics (for example, the **S** system) invalidate *Identity*, as do certain connexive logics (see, for example, E. P. Martin and R. K. Meyer, "Solution to the P–W Problem," *Journal of Symbolic Logic*, XLVII, 4 (December 1982): 869–87; Richard Sylvan, "A Preliminary Western History of Sociative Logics," in Dominic Hyde and Graham Priest, eds., *Sociative Logics and Their Applications: Essays by the Late Richard Sylvan* (Oxford: Ashgate, 2000); Robert K. Meyer and Errol P. Martin, "S (for Syllogism) Revisited: 'The Revolution Devours Its Children'," *Australasian Journal of Logic*, XVI, 3 (2019): 49–67; and Yale Weiss, "Connexive Extensions of Regular Conditional Logic," *Logic and Logical Philosophy*, XXVIII, 3 (September 2019): 611–27). However, these systems are not obviously meant to model the natural language conditional. For instance, Martin and Meyer, "Solution to the P–W problem," *op. cit.*, seem mainly interested in capturing intuitively valid, non-circular arguments; but

The one place where serious pressure has been put on *Identity* concerns its predictions about conditionals with logically inconsistent antecedents. *Identity* (in the presence of *Mon*) entails that conditionals with the form $\perp > p$ are all logically true. This corollary is accepted by some, but there is a serious case for rejecting it.²⁷ However, this particular instance of *Identity* does not play an essential role in my collapse result. We can distinguish *Identity* from a slightly weaker principle, *Identity* $_{\perp}$, which says that, whenever p is consistent, $\models p > p$. The proof in appendix B suffices to show that *Identity* $_{\perp}$ together with *Ad Falsum*, *Mon*, and *IE* leads to a slightly weakened collapse result: for any p and q , $p > q \models p \supset q$ provided that each of the following is (on its own) logically consistent: $\neg(p > q)$, $p \wedge q$, $\neg(p > \neg q)$, $p \wedge \neg q$. But these conditions obtain for almost all p and q which are logically orthogonal, and so this weakened result is still completely unacceptable.

So, even though *Identity* might well fail for conditionals with inconsistent antecedents, this does not help us evade my collapse result, which still follows, in an only slightly restricted form, from *Identity* $_{\perp}$. If you think that *Identity* $_{\perp}$ is plausible but *Identity* is not, then you can substitute *Identity* $_{\perp}$ for *Identity* throughout the paper, and almost everything I say will still apply.²⁸

Even in the absence of any theoretical case against *Identity*, we should, for the sake of completeness, still explore natural language

there is no reason to think that all logically true conditionals must represent intuitively non-circular arguments. Some trivalent systems invalidate *Identity* in a limited way on some notions of consequence, but not in ways that seem relevant for present purposes. Thanks to Yale Weiss for helpful discussion.

²⁷ See, for example, Matthias Jenny, “Counterpossibles in Science: The Case of Relative Computability,” *Noûs*, LII, 3 (September 2018): 530–60, and citations therein.

²⁸ Thanks to an anonymous referee for pushing me to address this point. It is important to distinguish *Logical Implication*—the conjunction of *Mon* and *Identity*—from some related principles which appear similar but are much less defensible. First, compare *Multi-Premise LI*, which says that, if $\Gamma, p \models q$, then $\Gamma \models p > q$. *Multi-Premise LI* is clearly false, for from *Multi-Premise LI* alone we can derive the conclusion that $p \supset q \models p > q$ (see Daniel Bonevac, Josh Dever, and David Sosa, “Conditionals and Their Antecedents,” unpublished manuscript (University of Texas-Austin, 2013)). Despite its superficial similarity to *Logical Implication*, however, it is easy to find intuitive grounds for rejecting *Multi-Premise LI*: 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 this 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. So there are very natural reasons to reject *Multi-Premise LI*; 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, compare the converse of *Logical Implication*, which says that, if $\models p > q$, then $p \models q$. I am inclined to accept *Converse LI*, and it is valid on the theory I adopt below. But its dialectical status is insecure, since it is in obvious tension with *IE*.

conditionals where it might plausibly fail if *IE* is valid. In particular, *Identity* says that sentences with the form $(q \wedge \neg(p > q)) > (q \wedge \neg(p > q))$ and hence (given *Mon*) $(q \wedge \neg(p > q)) > \neg(p > q)$ will invariably be true.²⁹ By contrast, *IE*-validating theories generally predict that these will invariably be false, provided the antecedents are possible.³⁰ This is not true of all theories, of course, in particular, of the material analysis; but it is true of most *IE*-validating theories, and hence these sentences seem like the most natural place to look for potential failures of *Identity*. So consider conditionals with this form, as in (7):

- (7) a. If the vase had broken, but it is not the case that the vase would have broken if it had been wrapped in plastic, then it is not the case that the vase would have broken if it had been wrapped in plastic.
 b. If the match had lit, but it is not the case that the match would have lit if it had been wet, then it is not the case that the match would have lit if it had been wet.

These feel like logical truths. So we do not find failures of *Identity* even where we might expect to find them from the point of view of *IE*-validating theories.

Note: not only do *IE*-validating theories generally *fail* to predict that sentences like (7) are logical truths; they also generally predict that their internal negations *are* logical truths.³¹ We can test that contrasting prediction by looking at a sentence like (8):

- (8) If the vase had broken, but it is 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.

(8) sounds trivially false to me, not at all like a logical truth.

Matters are similar for indicatives with the form $(q \wedge \neg(p > q)) > \neg(p > q)$:

- (9) a. If the vase broke, but it is not the case that the vase broke if it was wrapped in plastic, then it is not the case that the vase broke if it was wrapped in plastic.
 b. If the match lit, but it is not the case that it lit if it was wet, then it is not the case that it lit if it was wet.

²⁹In the proof in the second appendix, I look at sentences with conjunctive antecedents in the reverse order, but the present order seems somewhat smoother in English.

³⁰In the second appendix, we show that $(q \wedge \neg(p > q)) > (p > q)$ is a logical truth, given *IE* and our background assumptions. But then, by *Weak Conditional Non-Contradiction*, we can conclude $(q \wedge \neg(p > q)) > \neg(p > q)$ is false.

³¹As a referee for this JOURNAL helpfully points out, this reasoning would be blocked if we had only *Identity*_⊥, together with *Strong Centering* (which entails that $p \wedge q$ entails $p > q$), since in that case, $(q \wedge \neg(p > q)) \wedge p$ would be inconsistent.

Again, these feel like logical truths—in line with the predictions of *Identity*. (There is an interesting further element here, which is that both the conditionals in (9) strike me as somewhat odd; I return to this point in section VI.)

Intuitions about natural language thus do not give us any grounds for rejecting *Identity*. I conclude that we need a theory of the conditional that validates *Identity*.

Let me emphasize where that leaves us. A natural response to the collapse result that I presented above is to try to find a clever way to *block* the conclusion. For instance, we could adopt any of a variety of dynamic or informational notions of entailment, which would let us block certain moves in the proof. But this does not get us what we need. The collapse result is just a heuristic to help us understand *why* theories that validate *IE* tend to invalidate *Identity*. But what we need, in the end—what I take to have been obvious from the outset, but hopefully is even clearer now—is not a clever way of blocking the collapse result, but rather, simply, a theory which *does* validate *Identity*.³²

III.4. Import-Export. If we want to validate *Ad Falsum*, *Mon*, and *Identity*, and we do not want to collapse to the material conditional, then, given our result above, we must reject *IE*.

I think this conclusion is right. But this argument is indirect. It would be nice to either find 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.³³ Another, more abstract motivation for *IE* comes from Ramsey's famous suggestion that you should believe $p > q$ iff you believe q after adding p hypothetically to your stock of beliefs.³⁴ Repeated application of this test suggests that

³²Whereas dynamic theories like those of Paul Dekker, "Transsentential Meditations," *op. cit.*, p. 202; and Gillies, "On Truth Conditions for *If* (But Not Quite Only *If*)," *op. cit.*, do not, for the reasons we explored above.

³³For experimental evidence that confirms these intuitions, see Janneke van Wijnbergen-Huitink, Shira Elqayam, and David E. Over, "The Probability of Iterated Conditionals," *Cognitive Science*, xxxix, 4 (May 2015): 1–16.

³⁴Frank P. Ramsey, "General Propositions and Causality," in D. H. Mellor, ed., *Foundations: Essays in Philosophy, Logic, Mathematics and Economics* (London: Routledge and Kegan Paul, [1931] 1978), pp. 133–51.

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.³⁵

In the case of subjunctive conditionals, however, *IE* seems to break down. For instance, (10) (from David Etlin³⁶) and (11) (from Stephen Yablo, p.c.) instantiate the *IE* schema, but are not intuitively pairwise equivalent:

- (10) 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.
- (11) a. If I had been exactly 6' tall, then if I had been a bit taller than 6', I would have been 6'1".
 b. #If I had been exactly 6' tall and a bit taller than 6', I would have been 6'1".

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 (12):

- (12) 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.

(12-b) is obviously true, whereas we can certainly imagine (12-a) being false, if the strike would have prevented the exams being marked. For another example, suppose we have a die which is either weighted toward evens or odds; we do not know which. The die is never thrown. Compare (13-a) and (13-b):

- (13) a. If the die had been thrown and landed four, then if it had not landed four it would have landed two or six.
 b. #If the die had been thrown and landed four and it had not landed four, it would have landed two or six.

(13-a) is plausibly likely (but not certain) to be true: intuitively, it says that if the die had landed four, that would have shown that it was weighted toward evens, and so had it not landed four, it would have landed on an even still. Whereas (13-b) sounds incoherent.

Taken together, these examples provide compelling direct evidence which matches our indirect evidence: *IE* is not valid for subjunctives.

³⁵ See Arló-Costa, "Bayesian Epistemology and Epistemic Conditionals," *op. cit.* The argument is subtle and no doubt questionable, but suggestive.

³⁶ David Etlin, "Modus Ponens Revisited," unpublished manuscript (MIT, 2008).

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

- (14) a. If the match lit, then it lit if it was wet.
- b. If the match lit and it was wet, then it lit.
- (15) a. #If I am exactly 6' tall, then if I am a bit taller than 6', then I am 6'1".
- b. #If I am exactly 6' tall and a bit taller than 6', then I am 6'1".
- (16) 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.
- (17) a. #If the die was thrown and landed four, then if it did not land four it landed two or six.
- b. #If the die was thrown and landed four and it did not land four, it landed two or six.

Unlike the corresponding subjunctive pairs, these indicative versions appear to be pairwise equivalent. The lack of a counterexample in these cases of course does not necessarily show that *IE* is valid for indicatives. But since the very same pairs in the subjunctive mood strike us as inequivalent, this is at least suggestive that we will not find direct evidence against *IE* for indicatives. That is, *IE feels* valid for indicatives.³⁷

And this poses a real puzzle. We have powerful, but indirect, evidence that *IE* is not valid for indicative conditionals. But we do not 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.³⁸

IV. 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 is that an inference pattern may fail to preserve truth in *all* intended models, but still preserves truth

³⁷ Compare Richard Bradley, *Decision Theory with a Human Face* (Cambridge, UK: Cambridge University Press, 2017), p. 71, who calls *IE* "the defining characteristic of indicative conditionals"; Dale, "A Natural Deduction System for 'If Then'," *op. cit.*, too, endorses exportation for indicatives but not subjunctives.

³⁸ An anonymous referee for this JOURNAL points out that *IE* seems to fail for indicative unconditionals. 'If I'll get cancer, I'll get cancer whether or not I smoke' seems false; while the corresponding conjunction 'If I'll get cancer, I'll get cancer if I smoke, and if I'll get cancer, I'll get cancer if I do not smoke' seems true. More work is required to explore intuitions about *IE* in unconditionals.

in all cases *where we might plausibly use the sentences in question*. Then the inference pattern will not be logically valid, but it will be hard to find concrete counterexamples to it.

This is roughly the notion that von Fintel, following Strawson, spelled out as *Strawson validity*.³⁹ Here is my take on the notion:⁴⁰

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

A context, for us, is just a set of worlds which models a conversation's common ground.⁴¹ If an inference is Strawson valid, it does not necessarily preserve truth in all worlds in all models. But it does preserve truth in any context where all the premises and the conclusion have their presuppositions satisfied. This will plausibly include 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.

For a simple example of Strawson entailment, consider gender and number features on pronouns, which are plausibly a certain kind of presupposition.⁴² Hence the sentence 'She is female' plausibly presupposes that there is a salient referent for 'she' which is singular and female. When those requirements are fulfilled, the sentence in question will invariably be true. There is thus *some* sense in which this sentence is valid. On the other hand, the sentence is intuitively not *logically valid* ("I heard Sue got a cat." "Yeah, she's female." "No, it's a male cat."). So the validity of this sentence is well modeled as a Strawson validity: true on any ordinary or felicitous occasion of use, even if not always true.

Strawson validity is just one of a class of what we might call *Strawson concepts*: concepts that stipulate that certain relations hold between

³⁹ See Kai von Fintel, "NPI Licensing, Strawson Entailment, and Context Dependency," *Journal of Semantics*, xvi, 2 (1999): 97–148; and Peter Strawson, *Introduction to Logical Theory* (London: Methuen, 1952).

⁴⁰ The main difference from von Fintel's version is the relativization to a context, which plays a crucial role in formulating some presuppositional constraints, including the one I discuss below.

⁴¹ See Robert Stalnaker, "Pragmatic Presuppositions," in Milton K. Munitz and Peter Unger, eds., *Semantics and Philosophy* (New York: New York University Press, 1974), pp. 197–213; and Robert Stalnaker, "Assertion," in Peter Cole, ed., *Syntax and Semantics, Volume 9* (New York: Academic Press, 1978), pp. 315–22. The details of the model do not matter for our purposes; for simplicity I will often talk about the context worlds as the "epistemically accessible" ones.

⁴² For extensive discussion, see Yasutada Sudo, *On the Semantics of Phi Features on Pronouns*, PhD diss. (MIT, 2012).

sentences whenever their presuppositions are all satisfied. So, among others, we can also spell out a notion of *Strawson information validity*, which says that, in any context where the presuppositions of p and q are satisfied, if you accept p , then you accept q .⁴³

With these concepts 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. That is, I will propose that *Identity*, *Ad Falsum*, and *Mon* are logically valid, while *IE* is not logically valid, but is Strawson (informationally) valid for indicatives, but not subjunctives.⁴⁴

There is some controversy about what Strawson concepts tell us in general.⁴⁵ I think that some of those worries are compelling. But I will be putting Strawson entailment to a very limited use here: namely, filtering *acceptable* sentences from *unacceptable* ones. The idea will be that indicative conditionals have a certain presupposition, which helps guide us toward *IE*-validating and away from *IE*-invalidating interpretations. Interpretations which render *IE* invalid will always involve presupposition failures, and thus will be difficult to access intuitively. I think this relatively limited application of Strawson concepts is on good theoretical standing, whatever the status of Strawson concepts more generally speaking.

⁴³ More formally: if the presuppositions of both p and q are satisfied throughout $\{ \langle c, w \rangle : w \in c \}$, then, if p is true throughout c , then q is true throughout c . Cf. the very similar notion of *Strawsonian support-preserving consequence* in Justin Bledin, "Fatalism and the Logic of Unconditionals," *Noûs*, LIV, 1 (March 2020): 126–61.

⁴⁴ The kind of presupposition I will focus on here is plausibly very different from semantic presupposition, and may well be a different kind of thing altogether; in other work I use a different name for this dimension of meaning, 'bounds'. 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, makes it hard to distinguish sentences which are valid *because of their truth conditions* from those which are valid *because of their presuppositions* (cf. 'Either she is a cat or she is not a cat', versus 'She is female'; or in our framework, *Identity*, which I want to say is logically valid and differs from *IE*, which I want to say is only Strawson valid for indicatives). We can better capture them in a multi-dimensional approach (Hans G. Herzberger, "Dimensions of Truth," *Journal of Philosophical Logic*, II, 4 (October 1973): 535–56; Lauri Karttunen and Stanley Peters, "Conventional Implicatures in Montague Grammar," in C.-K. Oh and D. Dineen, eds., *Syntax and Semantics 11: Presupposition* (New York: Academic Press, 1979), pp. 1–56; Sudo, *On the Semantics of Phi Features on Pronouns*, *op. cit.*; and Cian Dorr and John Hawthorne, "If...: A Theory of Conditionals," unpublished manuscript (NYU and USC, 2018)), which is the approach I will follow. Formally, we can treat sentence meanings as functions which take a context and a world to a pair of a truth value and a presupposition value. My main points could, however, be recast in a trivalent setting. Thanks to Cian Dorr for helpful discussion on this point.

⁴⁵ See, in particular, Dorr and Hawthorne, "If...," *op. cit.*

V. THE INDICATIVE CONSTRAINT

In the rest of this paper, I will argue for a particular implementation of the idea that *IE* is Strawson valid for indicatives, while our other principles are logically valid. In this section, I will develop an independently motivated account of the differences between indicatives and subjunctives. Then I will show how this account helps with *IE*.

Following much of the literature on conditionals, going back to Stalnaker,⁴⁶ I will assume that indicative and subjunctive conditionals have, structurally speaking, the same truth conditions: they differ only in what selection function we use to evaluate them and (relatedly) in their presuppositions. For concreteness, I will build my proposal on top of Stalnakerian semantics,⁴⁷ though I should note that the general idea here could be accounted for in other frameworks, like Lewis's,⁴⁸ provided that they logically validate *Identity*, *Mon*, and *Ad Falsum*, and logically invalidate *IE*. On Stalnaker's theory, a conditional $p > q$ is true iff the closest p -world is a q -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.⁴⁹ Given selection functions f_i and f_s , where $>_i$ is the indicative conditional, $p >_i q$ is true at w iff q is true at $f_i(p, w)$; likewise, where $>_s$ is the subjunctive conditional, $p >_s q$ is true at w iff q is true at $f_s(p, w)$.

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, put forward by Stalnaker, and further developed, most importantly, by von Fintel.⁵⁰ The first says that indicative conditionals are always evaluated relative to a

⁴⁶ Robert Stalnaker, "Indicative Conditionals," *Philosophia*, v, 3 (July 1975): 269–86.

⁴⁷ Stalnaker, "A Theory of Conditionals," *op. cit.*; and Robert C. Stalnaker and Richmond H. Thomason, "A Semantic Analysis of Conditional Logic," *Theoria*, xxxvi, 1 (December 1970): 23–42.

⁴⁸ David Lewis, *Counterfactuals* (Oxford: Blackwell, 1973).

⁴⁹ Selection functions must satisfy the following conditions (we assume context provides a reflexive accessibility relation on worlds):

- *Strong Centering*: $f(p, w) = w$ iff $w \in p$;
- *Success*: $f(p, w) \in p$ provided p is true in a world accessible from w ;
- *CSO*: if $f(p, w) \in q$ and $f(q, w) \in p$, then $f(p, w) = f(q, w)$; and
- *Absurdity*: Where λ is an absurd world that makes all propositions true, $f(p, w) = \lambda$ iff p is true in no world accessible from w .

For readability, I will often use italics for both sentences and the corresponding propositions, ignoring relativization to contexts.

⁵⁰ Stalnaker, "Indicative conditionals," *op. cit.*; and Kai von Fintel, "The Presupposition of Subjunctive Conditionals," in O. Percus and U. Sauerland, eds., *The Interpretive Tract* (Cambridge, MA: MIT Working Papers in Linguistics, 1998), pp. 29–44.

selection function which treats contextually possible worlds as being closer to each other than any other worlds: that is, 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 .⁵¹ 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 $\neg p >_i q$ (this is the ‘or-to-‘if’ inference), but not necessarily the corresponding subjunctive $\neg p >_s q$. So, for instance, once we accept (18-a), then it seems we must also accept the indicative conditional (18-b), but not necessarily the subjunctive (18-c):

- (18) 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 closeness constraint is exactly what is needed, in the context of his theory, to account for this inference pattern. If Stalnaker’s constraint is satisfied, and $p \vee q$ is true throughout c , while $\neg p$ is compatible with c , then $\neg p >_i q$ will be true throughout c . That is because for any world in c , the closest $\neg p$ -world to that world, according to f_i , will be in c , hence will be in $p \vee q$, hence will be in q . Conversely, if a model does not satisfy Stalnaker’s closeness constraint, we will be able to construct failures of ‘or-to-‘if’ in that model.

The second proposal says that indicatives presuppose that their antecedents are epistemically possible: $p >_i q$ is felicitous only in a context compatible with p , whereas $p >_s q$ can be felicitous even in a context incompatible with p .⁵² 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:

⁵¹ Compare a very similar proposal in W. L. Harper, “Ramsey Test Conditionals and Iterated Belief Change (A Response to Stalnaker),” in W. L. Harper and C. A. Hooker, eds., *Foundations of Probability Theory, Statistical Inference, and Statistical Theories of Science, Volume I* (Dordrecht, the Netherlands: D. Reidel, 1976), pp. 117–35.

⁵² In addition to Stalnaker and von Fintel, see also Gillies, “On Truth Conditions for *If* (But Not Quite Only *If*),” *op. cit.*; Brian Leahy, “Presuppositions and Antipresuppositions in Conditionals,” in Neil Ashton, Anca Chereches, and David Lutz, eds., *Semantics and Linguistic Theory (SALT)*, vol. 21 (New Brunswick, NJ: Linguistic Society of America and Rutgers University, 2011), pp. 257–74; John Mackay, “Modal Interpretation of Tense in Subjunctive Conditionals,” *Semantics and Pragmatics*, xii, 2 (2019): 1–29; and Ben Holguín, “Knowledge in the Face of Conspiracy Conditionals,” *Linguistics and Philosophy*, xliiv (2021): 737–71, for more recent discussion. Kevin Dorst, “Abominable KK Failures,” *Mind*, cxxviii, 512 (October 2019): 1227–59, argues against this compatibility constraint. His data, however, seem to be order sensitive in a way that suggests that they involve a context shift; see Holguín, “Knowledge in the Face of Conspiracy Conditionals,” *op. cit.*, for discussion to this effect.

- (19) John didn't come to the party.
 a. #If he came, it was a disaster.
 b. If he had come, it would have been a disaster.

Both of these constraints—which are independent from each other—are well motivated. To capture them both in our semantic theory, we say that indicative conditionals presuppose that the selected antecedent world from any context world is itself in the context. That is, $p >_i q$ presupposes, at a context c , that, $\forall w \in c : f_i(p, w) \in c$. I will call this the *indicative constraint*.⁵³

So far, this is the standard Stalnakerian story. I want to propose a variation on this picture, on which the indicative presupposition is implemented in a *local* way.⁵⁴ To see this point, consider the contrast in (20):

- (20) 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 (20-a) is infelicitous, in contrast to the subjunctive variant in (20-b). But this is surprising from the point of view of the standard picture just sketched, because, relative to the global context in (20), it is epistemically possible that Bob did not go to the party, and so, globally speaking, the compatibility part of the indicative constraint seems to be satisfied. To account for the contrast in (20), 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 (20-a)—that Bob came to the party.

Similar points can be made in a variety of other environments. For instance, consider the pair of quantified conditionals in (21):

⁵³The closeness constraint immediately follows from this formulation. The compatibility constraint does as well, since by the structure of selection functions, $f_i(p, w)$ must be in p unless p is true in no accessible world, in which case it will be λ ; but since λ is (by assumption) 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.

⁵⁴Thanks to Cian Dorr, Irene Heim, and Ginger Schultheis for very helpful discussion of this idea. David Boylan and Ginger Schultheis, "The Qualitative Thesis," this JOURNAL, forthcoming (2020), independently propose a similar but subtly different account, motivated on different grounds. Their constraint captures all the locality data I have discussed here, but it does not validate *Identity*, because selection functions in their theory are assumed to shift with local contexts. This 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.

- (21) 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 (21), for every student x , we leave it open that x studied. Yet the indicative conditional in (21-a), 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 (21-b) 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 (an instance especially pertinent to our broader interests here) comes from nested conditionals. Suppose, again, that we have a die which is either weighted toward evens or odds; we do not know which, and we do not know whether the die was thrown. Compare again:

- (22) 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 in (22-a) and (22-b)—that the die did not 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 (22-b) seems acceptable, while the indicative variant is not. (22-b) has a clear meaning: it communicates that, if the die had landed four, then it would have been weighted toward evens, and so would have landed two or six if not four. In the scenario, this is likely true. By contrast, the indicative variant in (22-a) just sounds incoherent. 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. Hence consider (23):

- (23) 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 accessibility for epistemic modals is typically calculated in a local manner in general.⁵⁵ This can be illustrated by variants on the conditional cases we have just looked at, but which replace the indicative ‘If p, then q’ with ‘Might p’:

- (24) 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 will study. But every student who doesn’t study might study.
 c. #We don’t know whether the die will be thrown. But if the die will be thrown and will land four, then it might not land four.

In each of these cases the prejacent of the ‘might’ is compatible with what is globally epistemically possible; nonetheless, the sentences are infelicitous. We might expect a sentence like the second sentence in (24-a) to mean the same as ‘Suppose that Bob came to the party, and that, for all we know, he didn’t come’; but since the latter is perfectly coherent, these apparently differ in meaning, as Yalcin argues.⁵⁶ Cases like these suggest that epistemic possibility in general is calculated locally, too. Given this, it is not surprising that the indicative constraint should also be calculated locally.

Unsurprising, but not trivial. You might think that this simply follows from the “global” indicative constraint when we couple it with standard theories of presupposition projection, which say that a clause’s presuppositions must be satisfied throughout its local context. But that is wrong. If we simply took on board the global indicative constraint, and said that it must be satisfied throughout its local context, this would not account for these data. For the global indicative constraint to be satisfied throughout a local context, we would

⁵⁵ Especially the dynamic approach, for which see Jeroen Groenendijk, Martin Stokhof, and Frank Veltman, “Coreference and Modality,” in Shalom Lappin, ed., *Handbook of Contemporary Semantic Theory* (Oxford: Blackwell, 1996), pp. 179–216; Maria D. Aloni, *Quantification under Conceptual Covers*, PhD diss. (University of Amsterdam, 2001); and Seth Yalcin, “Epistemic Modality *De Re*,” *Ergo*, 11, 19 (2015): 475–527, but also on some pragmatic approaches like Cian Dorr and John Hawthorne, “Embedding Epistemic Modals,” *Mind*, cxxii, 488 (October 2013): 867–913, as well as my approach in Matthew Mandelkern, “Bounded Modality,” *The Philosophical Review*, cxxviii, 1 (January 2019): 1–61.

⁵⁶ Yalcin, “Epistemic Modals,” *op. cit.*

need that, for every world in the local context for the conditional, the global indicative constraint is satisfied at that world. But the global indicative constraint already quantifies over worlds, so the quantification over local context worlds is trivial here: this is just equivalent to the global indicative constraint.⁵⁷ We thus need a different way of tying the quantification in the indicative constraint directly to its local context.

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,⁵⁸ the domain theory,⁵⁹ the salience-based theory,⁶⁰ or the bounded theory.⁶¹ Here I will build on the bounded theory. The choice is largely because I think that is the right theory of epistemic modality, for the reasons I discuss elsewhere,⁶² most of the arguments I make there carry over to conditionals. I will not justify that choice at length here, 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 reiterate, however, why I am not building on a dynamic theory, which would in some ways be the most obvious approach. This is because, on the dynamic theory, the interpretation of embedded modals and conditionals is sensitive to local information in such a way that principles like *Identity* end up being invalid. By contrast, in the theory I develop here, local information does not *shift* the interpretation of embedded conditionals, but rather *bounds* the range of possible meanings by way of presuppositions. This is a crucial distinction, at the heart of how we will Strawson validate *IE* without invalidating *Identity*.

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.⁶³ A local context is a set of worlds which represents the information locally

⁵⁷ More formally, if κ is the conditional’s local context, and c is the global context, the global indicative constraint is satisfied throughout κ iff $\forall w \in \kappa : \forall w' \in c : f_i(p, w') \in c$. But the first layer of quantification here is vacuous.

⁵⁸ Groenendijk, Stokhof, and Veltman, “Coreference and Modality,” *op. cit.*; and Aloni, *Quantification under Conceptual Covers*, *op. cit.*

⁵⁹ Yalcin, “Epistemic Modals,” *op. cit.*; MacFarlane, “Epistemic Modals Are Assessment Sensitive,” *op. cit.*; and Nathan Klinedinst and Daniel Rothschild, “Connectives without Truthables,” *Natural Language Semantics*, xx, 2 (June 2012): 137–75.

⁶⁰ Dorr and Hawthorne, “Embedding Epistemic Modals,” *op. cit.*

⁶¹ Mandelkern, “Bounded Modality,” *op. cit.*

⁶² *Ibid.*

⁶³ See Philippe Schlenker, “Be Articulate: A Pragmatic Theory of Presupposition Projection,” *Theoretical Linguistics*, xxxiv, 3 (December 2008): 157–212; and Philippe

available relative to a given syntactic environment 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. This suffices to account for embedding data like those we saw briefly above.

I propose to implement a local version of the indicative constraint on analogy to the bounded theory's locality presupposition. In fact, we do not 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' to get the desired local version. In other words, where κ is the conditional's local context, our *locality constraint* says that $p >_i q$ presupposes that $\forall w \in \kappa : f_i(p, w) \in \kappa$.

How does this answer to the motivations given above? For unembedded conditionals, the locality constraint is equivalent to the standard indicative constraint. 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 $c^{\neg p}$. (In general, I will write c^p for the set of worlds in c where p is true and has its presuppositions satisfied, relative to c .) That means that, if the locality constraint is satisfied, then for any world w' in $c^{\neg p}$, $f_i(p, w') \in c^{\neg p}$. 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 locality constraint.

Schlenker, "Local Contexts," *Semantics and Pragmatics*, II, 3 (2009): 1–78. 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, embedded 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 (25):

- (25) 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.

Crucially, this reasoning goes through whether $\neg p \wedge (p >_i q)$ is embedded or unembedded, accounting for the infelicity of sentences which embed $\neg p \wedge (p >_i q)$, like (20-a).

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 (21-a), 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 entails the antecedent. So, in a conditional with the form $p >_i (\neg p >_i q)$, the local context for the consequent is c^p , and so the locality constraint for the embedded conditional will entail that, for any world $w' \in c^p$, $f_i(\neg p, w') \in c^p$. This will, however, again be impossible, since there are no $\neg p$ -worlds in c^p , thus accounting for the infelicity of sentences like (17-a). Since—I will assume—subjunctive conditionals do not have a corresponding locality constraint, none of this reasoning will go through for subjunctives, accounting for the contrasts observed in (13)–(20). This informal exposition will suffice for our purposes, but I summarize the resulting picture more carefully in a footnote.⁶⁴

VI. BACK TO IMPORT-EXPORT

I will now return to the questions about the logic of conditionals which started us off. The locality constraint—which we have so far motivated with observations about embedded conditionals—has surprising and desirable consequences for logic: it entails that *IE* is Strawson (informationally) valid for indicatives, but not subjunctives.

Start with Strawson informational equivalence. Recall that the pairs that instantiate *IE* have the form of (26-b) and (26-a), respectively:

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

⁶⁴I use 'satt' as shorthand for 'has its presuppositions satisfied'. $p \wedge q$ is satt, relative to κ , iff p is satt relative to κ^p and q is satt relative to κ^q . $p \vee q$ is satt, relative to κ , iff p is satt relative to $\kappa^{\neg p}$ and q is satt relative to $\kappa^{\neg q}$. And $\neg p$ is satt relative to κ iff p is satt relative to κ . We can treat the material conditional $p \supset q$ as equivalent to $\neg p \vee q$, and the material biconditional as equivalent to $(p \supset q) \wedge (q \supset p)$. The truth conditions for all these are classical.

Conditionals have Stalnaker's truth conditions. As for presuppositions, $p >_i q$ is satt relative to κ iff (i) p is satt, relative to κ ; (ii) q is satt, relative to κ^p ; and (iii) $\kappa \neq \emptyset \wedge \forall w \in \kappa : f_i(p, w) \in \kappa$. And $p >_s q$ is satt relative to κ iff p is satt relative to $\bigcup_{w \in \kappa} R(w)$, the set of all worlds accessible from a world in κ , and q is satt relative to $\bigcup_{w \in \kappa} f_s(p, w)$.

A different approach to subjunctives, suggested by Ginger Schultheis, "Counterfactual Probability," unpublished manuscript (University of Chicago, 2020), would ascribe locality to both indicatives and subjunctives, but let the role of the subjunctive mood be to expand local contexts. I like Schultheis's proposal, but for reasons that go beyond present concerns.

Consider a context s , and suppose that the locality constraints of (26-a) and (26-b) are satisfied throughout s . Then (26-a) is accepted in s (that is, true throughout s) just in case (26-b) is. Thus, in the terminology introduced above, (26-a) and (26-b) are Strawson informationally equivalent.

The proof turns on the observation that, provided the locality constraints of the conditionals are satisfied throughout the set of worlds compatible with what you accept, you accept either conditional just in case all the $p \wedge q$ -worlds compatible with what you accept are r -worlds. To see this, let s be the set of worlds compatible with what you accept. Suppose first that s contains a $p \wedge q$ -world which is not an r -world; then both (26-a) and (26-b) will obviously each be false at that world. Suppose next that every $p \wedge q$ -world in s is an r -world. Consider an arbitrary world w in s . (26-a) is true at w iff the closest $p \wedge q$ -world to w , call it w_{pq} , is an r -world; by the locality constraint, w_{pq} must be an s -world; but since all $p \wedge q$ -worlds in s are r -worlds by assumption, w_{pq} is an r -world and so (26-a) is true at w . The *locality* part of the locality constraint did not play a crucial role here, but it does play a crucial role in our reasoning about (26-b). (26-b) is true at w iff the closest p -world to w , call it w_p , is such that the closest q -world to w_p , call it w_{pq} , is an r -world. The locality constraint of the embedded conditional $q >_i r$ ensures that w_{pq} is in that conditional's local context, that is, s^p . So, w_{pq} is a $p \wedge q$ -world in s , and thus is an r -world by assumption; so (26-b) is true at w . So, provided their presuppositions are satisfied throughout a set of worlds, (26-a) and (26-b) are true throughout that set of worlds under exactly the same circumstances. This reasoning turns crucially on the locality constraint, so nothing similar follows for the subjunctive analogs of (26-a) and (26-b).

Very similar reasoning shows that *IE* is Strawson valid for indicatives. I leave the proof of this in a footnote.⁶⁵

⁶⁵ *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 $\langle c, w \rangle$ but this material conditional is false at $\langle c, w \rangle$. 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 $\langle c^{p >_i (q >_i r)}, w \rangle$. By the locality constraint, since $w \in c^{p >_i (q >_i r)}$, $f_i(\llbracket p \wedge q \rrbracket^c, w) \in c^{p >_i (q >_i r)}$. 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 $f_i(\llbracket p \wedge q \rrbracket^c, w)$, 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 $\langle c, w \rangle$ but the material conditional is false at $\langle c, w \rangle$. Then $p >_i (q >_i r)$

Taken together, these points predict a kind of indistinguishability for indicative 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 will not 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. This is because these are Strawson informationally equivalent. You might still worry that, as long as it is possible for one of these sentences to be true and the other to be false, 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 true *and* have their presuppositions satisfied.

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 *constraint* on the conditional's domain of quantification. In other words, it bounds the possible intended 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 *Identity*, *Mon*, and *Ad Falsum*, and also avoid the collapse result above: our conditional is not the material conditional. All these points are immediate from the fact that they

r) is false at $\langle c^{(p \wedge q) >_i r}, w \rangle$. By the locality constraints, $f_i(\llbracket p \rrbracket^c, w) \in c^{(p \wedge q) >_i r}$; and so again by the locality constraints, $f_i(\llbracket q \rrbracket^c, f_i(\llbracket p \rrbracket^c, w)) \in c^{((p \wedge q) >_i r) \wedge p}$, 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 $\langle c^{(p \wedge q) >_i r}, w \rangle$ after all, contrary to assumption.

Note that, because of the way ' \supset ' manipulates local contexts, we do not have a converse 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. 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. Or perhaps not. An alternative approach would be to marry the locality constraint with the theory of conditionals in van Fraassen, "Probabilities of Conditionals," *op. cit.*, which is a slight strengthening of Stalnaker's theory. As Cian Dorr has pointed out, the result would Strawson validate both meta-language and object-language *IE*. I am sympathetic to this latter approach, which I explore in work in progress.

hold of Stalnaker's logic, and our logic is just exactly Stalnaker's. You might worry that this ignores local contexts. But this is legitimate, because local contexts affect presupposition satisfaction but never truth in our system. That means that when we look at the logic of the system, we can simply ignore them. This is a formally convenient fact which means that our conditional has exactly the logic of Stalnaker's conditional, since it differs from Stalnaker's conditional only with respect to its presuppositions.

The Strawson logic of our conditional will be a strict strengthening of Stalnaker's logic. But the material conditional still does not Strawson entail the conditional; and so inferences like those from $\neg(p >_i q)$ to $p \wedge \neg q$, which are so disastrously valid on the material conditional, will still not be either logically or Strawson valid for our conditional.

While $p \supset q$ does not logically or Strawson entail $p >_i q$, $p \supset q$ does Strawson *informationally* entail $p >_i q$. So our approach walks a fine line. I think this is exactly the line we need to walk, however, since there is evidence that the inference from $p \supset q$ to $p >_i q$ is Strawson informationally valid: once you accept $p \supset q$, it seems like you are forced to accept the indicative conditional $p >_i q$, but not $p >_s q$, as we saw in section v. This is just the 'or'-to-'if' inference.⁶⁶

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 *Identity* together with *Mon*; whereas *IE* says that sentences with the form $(\neg(p >_i q) \wedge q) >_i (p >_i q)$ are instead valid. Our theory of the indicative validates *Identity* and not *IE*, and so predicts that sentences with the first form are always true, while sentences of the second form are never true when their antecedents are possible. 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 $(q \wedge \neg(p >_i q))$.⁶⁷ So our account predicts

⁶⁶For more exploration of the relationship between the indicative and the material conditional in a framework much like the present one, see Boylan and Schultheis, "The Qualitative Thesis," *op. cit.* An entirely separate case for the locality constraint can be made on the basis of the kinds of considerations they discuss, which involve a local version of the 'or'-to-'if' inference.

⁶⁷This 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.

that these two crucial premises in our collapse result will, in the case of indicative conditionals, never have their presuppositions satisfied. And this looks right: indicatives with this form, as we noted when we looked at (9-b), sound very strange. For another example, compare:

- (27) 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. . .

Contrasts like this suggest that our theory avoids the collapse 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 *Identity* 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 *Identity* is only Strawson valid, and *IE* logically valid. I want to note two responses. First, although our key conditionals in the indicative mood are indeed somewhat odd, I still think that intuition favors the predictions of *Identity* over those of *IE*, as I argued in section III.3. Second, as we have seen, in the subjunctive case things look different: there our target conditionals are felicitous, and the predictions of *Identity*, not *IE*, are clearly correct. So for subjunctives, it seems clear that we should logically validate *Identity*, and (in every sense) invalidate *IE*. Insofar as we want as unified as possible a theory of indicatives and subjunctives, then, we should logically validate *Identity* and logically invalidate *IE* for all conditionals, and then find a way to predict that the latter is Strawson valid for indicatives.

VII. CONCLUSION

An adequate theory of the conditional must navigate a narrow passage between, on the one hand, apparently inviolable logical principles; and, on the other, the obviously untenable material analysis. In this paper I have showed that this passage is even narrower than it appeared after the famous result of Dale/Gibbard. In particular, the only way to validate *Identity*, *IE*, *Mon*, and *Ad Falsum* together is with the material conditional. This result helps explain why extant theories of the conditional which validate *IE* invalidate *Identity*, and, given the plausibility of *Identity*, it amounts to a clear argument against *IE*.

But can we find direct evidence that *IE* is invalid? In the case of subjunctive conditionals, we can; but not, apparently, for indicatives. In the second part of the paper, I developed a theory which aims to account for these subtle facts. The truth conditions are Stalnaker's,

and so the logic is Stalnaker's. On top of those truth conditions, I proposed a local implementation of Stalnaker's indicative constraint, which I motivated on the basis of a range of contrasts in embedding behavior between indicatives and subjunctives. This constraint has a surprising result: even though *IE* is not logically valid for indicatives, it is Strawson (informationally) valid. This helps us make sense of the lack of apparent counterexamples to *IE* for indicatives in a logically hygienic way.

My goal here has been partly to advocate a positive account, but equally to try to make clear the desiderata for any account. We must, *inter alia*, validate *Identity*, and account for the contrast between indicatives and subjunctives with respect to *IE* and local compatibility. While I am sympathetic to the positive proposal I have sketched here, I am open to other ways of accounting for intuitions in this area; what I hope to have shown clearly is that it is *not* an option to simply straightforwardly semantically validate *IE* and thereby invalidate *Identity*.⁶⁸

There is much more to explore about the logic of indicative and subjunctive conditionals. For instance: How can we make sense of the fact that we tend to assign the same probabilities to indicative pairs that instantiate *IE*? How can we make sense of the fact that *IE* feels valid *as a default matter* for subjunctives, even if, as we have seen, it does not appear to be valid in general for subjunctives? How can we make sense of apparent counterexamples to *MP* of the kind discussed by McGee,⁶⁹ if we adopt a theory, like the one I have developed, that validates *MP*?

There is much work left to do here. But I am hopeful that the general strategy taken here—recruiting independently motivated theories about the contrasts between indicatives and subjunctives to account for their apparent logical differences—will prove fruitful. In

⁶⁸ One possibility is to validate *IE* on a restricted basis, for only Boolean antecedents, as in Yalcin, "Epistemic Modals," *op. cit.*; Justin Bledin, "Modus Ponens Defended," this JOURNAL, CXII, 2 (February 2015): 57–83; Ciardelli, "Indicative Conditionals and Graded Information," *op. cit.*; John Cantwell, "An Expressivist Analysis of the Indicative Conditional with a Restrictor Semantics," *Review of Symbolic Logic*, forthcoming (2020); and Norlin, "Acceptance, Certainty, and Indicative Conditionals," *op. cit.*; it is unclear, however, how this extends to subjunctives. Another approach would be to amend the restrictor theory so that it validates *Identity* (and thus, again, only a limited form of *IE*). In the standard theory, we have $\llbracket p > q \rrbracket^{f,g,w} = 1$ iff $\llbracket q \rrbracket^{f,g,w} = 1$, where f is a modal base and g an ordering source, with f^b defined as the smallest function which takes any world w to $f(w) \cup \{\llbracket p \rrbracket^{f,g}\}$. We could instead define $f^b(w)$ as the least (in a sense which would need to be clarified) expansion of $f(w)$ such that p is true at every world in $\bigcap f^b(w)$, assessed relative to f^b and g , if there is one, and undefined otherwise.

⁶⁹ McGee, "A Counterexample to Modus Ponens," *op. cit.*

particular, this strategy may help us better understand how the conditional can occupy the narrow space between, on the one hand, *prima facie* plausible principles like *Identity* and *IE*; and, on the other, unacceptable logical results like the collapse of the conditional to the material conditional.

APPENDIX A. FAILURES OF *IDENTITY* IN MCGEE'S THEORY

We show that *Identity* is not valid in McGee's theory,⁷⁰ by explaining his semantics and then showing something stronger: that there are sentences with the form $p > p$ which are *never* true on that semantics in any non-trivial model (that is, any model with at least one atom which is true at some world and false at another).

On McGee's theory, sentences are evaluated relative to two parameters.⁷¹ The first is a Stalnakerian *selection function* f from consistent propositions and worlds to worlds. The second is a set of sentences Γ , which keeps track of conditional antecedents. With \mathfrak{J} an atomic valuation function, A any atom, and p, q any sentences, we have:

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

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 . Our target instance of *Identity* is $(\neg(\neg A > A) \wedge A) > (\neg(\neg A > A) \wedge A)$. Assume for contradiction that our target sentence is true, relative to the empty set, at w : that is, assume $\llbracket (\neg(\neg A > A) \wedge A) > (\neg(\neg A > A) \wedge A) \rrbracket^{\emptyset, w} = 1$. Then we can reason as follows:

1. $\llbracket (\neg(\neg A > A) \wedge A) \rrbracket^{\{\neg(\neg A > A) \wedge A\}, w} = 1$ *By Cond*
2. $\llbracket \neg(\neg A > A) \rrbracket^{\{\neg(\neg A > A) \wedge A\}, w} = 1$ *By Conj*
3. $\llbracket \neg A > A \rrbracket^{\{\neg(\neg A > A) \wedge A\}, w} = 0$

By Neg, since $\neg(\neg A > A) \wedge A$ is consistent

⁷⁰ *Ibid.*

⁷¹ Actually there is a third—a binary reflexive accessibility relation—but that does not matter for the present purpose, so I ignore it. There will still be failures of *Identity* if we add the accessibility relation back in. (The semantics I give is equivalent to the case where accessibility is universal.)

$$4. \quad \llbracket A \rrbracket^{\{\neg(\neg A > A) \wedge A, \neg A\}, w} = 0$$

By *Cond*

But $\llbracket A \rrbracket^{\{\neg(\neg A > A) \wedge A, \neg A\}, w} = 1$ by *Absurd*, since $\llbracket \neg(\neg A > A) \wedge A \rrbracket^\emptyset \cap \llbracket \neg A \rrbracket^\emptyset = \emptyset$ by classical reasoning which remains valid in this setting.

So we have derived a contradiction from the assumption that $(\neg(\neg A > A) \wedge A) > (\neg(\neg A > A) \wedge A)$ is true. Since w and f were chosen arbitrarily, $(\neg(\neg A > A) \wedge A) > (\neg(\neg A > A) \wedge A)$ is false relative to the empty set at any world and selection function in any model that includes both A - and $\neg A$ -worlds.

APPENDIX B. *IDENTITY + MON + IE + AD FALSUM LEAD TO COLLAPSE*

We assume *Identity*, *Mon*, *IE*, and *Ad Falsum*; classical properties of entailment (\models) and conjunction, the material conditional, and negation; and substitutability of $p \wedge q$ for $q \wedge p$ in conditional antecedents. We show that for any a, c , it follows that $a > c \models a \supset c$.⁷²

First, note that, for any a, c , if $a \models c$, then $\models a > c$; call this principle *Logical Implication (LI)*. *LI* is immediate from *Identity* together with *Mon*, since by *Identity* we have $\models a > a$ and hence by *Mon* we have $\models a > c$. By the monotonicity of \models , it also follows that whenever $\models c$, we also have $a \models c$ and hence (by *LI*) $\models a > c$.

We start by proving $a > c \models a \supset c$:

- | | |
|--|---|
| 1. $\models \neg(a > \neg c) > \neg(a > \neg c)$ | <i>Identity</i> |
| 2. $\models \neg c > (\neg(a > \neg c) > \neg(a > \neg c))$ | <i>LI, 1</i> |
| 3. $\models (\neg c \wedge \neg(a > \neg c)) > \neg(a > \neg c)$ | <i>IE, 2</i> |
| 4. $\models (\neg(a > \neg c) \wedge \neg c) > \neg(a > \neg c)$ | <i>Substitution, 3</i> |
| 5. $\models (\neg c \wedge a) > \neg c$ | <i>LI, classical logic</i> |
| 6. $\models \neg c > (a > \neg c)$ | <i>IE, 5</i> |
| 7. $\models \neg(a > \neg c) > (\neg c > (a > \neg c))$ | <i>LI, 6</i> |
| 8. $\models (\neg(a > \neg c) \wedge \neg c) > (a > \neg c)$ | <i>IE, 7</i> |
| 9. $\models \neg(\neg(a > \neg c) \wedge \neg c)$ | <i>Ad Falsum, 4, 8</i> |
| 10. $\neg c \models a > \neg c$ | <i>Classical logic, 9</i> |
| 11. $(a > c) \wedge \neg c \models (a > c) \wedge (a > \neg c)$ | <i>Classical logic, 10</i> |
| 12. $(a > c) \wedge (a > \neg c) \models \neg a$ | <i>Ad Falsum</i> |
| 13. $(a > c) \wedge \neg c \models \neg a$ | <i>Transitivity of \models, 11, 12</i> |
| 14. $a > c \models a \supset c$ | <i>Classical logic, 13</i> |

Next we prove $a \supset c \models a > c$:

- | | |
|--|----------------------------|
| 1. $\models ((a \supset c) \wedge a) > c$ | <i>LI, Classical logic</i> |
| 2. $\models (a \supset c) > (a > c)$ | <i>IE, 1</i> |
| 3. $\models \neg(a > c) > ((a \supset c) > (a > c))$ | <i>LI, 2</i> |

⁷²This proof is not maximally concise, but the shorter versions I have found rely on instances of *Identity* with logically inconsistent antecedents, while this version does not.

- | | | |
|-----|--|---------------------------|
| 4. | $\models (\neg(a > c) \wedge (a \supset c)) > (a > c)$ | <i>IE, 3</i> |
| 5. | $\models ((a \supset c) \wedge \neg(a > c)) > (a > c)$ | <i>Substitution, 4</i> |
| 6. | $\models \neg(a > c) > \neg(a > c)$ | <i>Identity</i> |
| 7. | $\models (a \supset c) > (\neg(a > c) > \neg(a > c))$ | <i>LI, 6</i> |
| 8. | $\models ((a \supset c) \wedge \neg(a > c)) > \neg(a > c)$ | <i>IE, 7</i> |
| 9. | $\models \neg((a \supset c) \wedge \neg(a > c))$ | <i>Ad Falsum, 5, 8</i> |
| 10. | $a \supset c \models a > c$ | <i>Classical logic, 9</i> |

New York University

MATTHEW MANDELKERN