

Note added September 2010: I am no longer convinced by the arguments in this paper, and do not believe the conclusion here defended. However, this version has been cited by a couple of papers, so I will continue to make it available for those who are interested. (I also still think it makes a couple of useful points.)

*‘Might’ Counterfactuals**

Antony Eagle

17 September 2007

Abstract

A ‘might’ counterfactual is a sentence of the form ‘If it had been the case that *A*, it might have been the case that *C*’. Recently, John Hawthorne has argued that the truth of many ‘might’ counterfactuals precludes the truth of most ‘would’ counterfactuals. I examine the semantics of ‘might’ counterfactuals, with one eye towards defusing this argument, but mostly with the aim of understanding this interesting class of sentences better.

1 The Argument from Inescapable Clashes

Recently, John Hawthorne argued that, in a chancy quantum mechanical world,

[w]e shall certainly be tempted to think that most ordinary counterfactuals are false. After all, having assimilated the [quantum] theory, we shall be led to accept:

- (1) If I had dropped the plate, it might have flown off sideways.

This in turn will induce us to think that (2) is incorrect:

- (2) If I had dropped the plate, it would have fallen to the floor.

... we then conclude that those propositions expressed by ordinary counterfactuals like (2) are false. (Hawthorne, 2005: 396)

Following DeRose (1999), we can call the conclusion of the argument *counterfactual scepticism*. In Hawthorne’s case, we find the ‘might’ counterfactual (1) compelling on physical grounds, and recognise that there is an intuitive *clash* with a related ‘would’ counterfactual (2). Since the ‘might’ claim seems in some sense

*Thanks to Keith DeRose, Alan Hájek, Daniel Nolan, Robbie Williams, Tim Williamson and an audience at Leeds for helpful discussions.

weaker, and moreover there is a sentence that can conceivably adequately replace the ‘would’ claim and is compatible with the ‘might’ claim (viz. ‘If I had dropped the plate, it would very likely have fallen to the floor’), we are tempted here to reject (2). But (2) isn’t special; if the world is chancy in the way that a straightforward understanding of quantum mechanics tells us, very many sentences of the form ‘if it had been that p , it might have been that q ’ are true. In each case, there is a related would counterfactual that will be false. Even if there is the weakened replacement (with ‘likely’ modifying the consequent), that still means that almost every straightforward ordinary counterfactual claim we assent to is false, a quite unpalatable conclusion.

In general, we can call this the argument from *inescapable clashes*. Ordinary speakers find sentences of the form (3) quite unacceptable and clashing:

- (3) If it had been that p , it would have been that q ; and if it had been that p , it might not have been that q .

Ordinary speakers also recognise that the second conjunct is quite weak, and since they cannot accept both conjuncts, they accept the ‘might’ counterfactual and reject the ‘would’ counterfactual.

As Hawthorne continues, the argument from inescapable clashes drops rapidly from sight—he is more concerned that the similarity metric for ‘would’ counterfactuals defended by Lewis (1986a) already renders most ordinary counterfactuals false, regardless of whether we make this detour through ‘might’ counterfactuals. I will not be concerned here with this problem for Lewis’ system, or whether there are any repairs, broadly in the Humean tradition, that can save it (Williams, unpublished). For the argument from inescapable clashes is a deep threat to ordinary ‘would’ counterfactuals, no matter what one says about the details of their semantics.

The reason the argument is so pressing is that, because ordinary speakers find (3) so unacceptable, any semantic account of ‘might’ and ‘would’ counterfactuals should make (3) come out false. Any theory of ‘might’ and ‘would’ counterfactuals that makes (3) come out unproblematically true is *prima facie* defective (DeRose, 1999: 395). So if we wish to resist the argument from inescapable clashes, we must reject the step from the falsity of (3) to the falsity of the first conjunct of (3). But that step was motivated by the fact that the ‘might’ counterfactuals sound very weak, and difficult to dissent from, in a way that the ‘would’ counterfactual is (especially in the context of (3)) quite easy to dissent from. It’s difficult to see what

other responses are available. Moreover, in many cases (like quantum mechanics) we want to say things like ‘if p , there is some chance of q ’ (‘If I were to toss the coin, it has some chance of landing heads’); and it seems a simple step from this claim to the truth of the second conjunct of (3).¹

Difficult, perhaps, but not impossible. I take it to be quite plausible that the problems raised by the argument from inescapable clashes are to be resolved by giving a good account of the relatively underexplored ‘might’ counterfactuals, rather than messing fundamentally with the much better studied and known ‘would’ counterfactuals. In what follows, I aim to explore the options for an account of the ‘might’ counterfactual that will allow sentences like (3) to be true, while nevertheless explaining what makes them sound so awful to ordinary speakers. This will both save the phenomenon of inescapable clashes, while defeating the counterfactual scepticism that depends on such clashes.²

2 Introducing Theories of ‘Might’ Counterfactuals

Three main theories of the semantics of ‘might’ counterfactuals have found adherents in the literature. Before discussing them in detail, I want to present broad sketches of what they involve. The labels I use haven’t become standard, and the three positions lump together subtly different positions, yet this should provide a useful starting point:³

Duality The ‘might’ counterfactual ‘if it had been that p , it might have been that q ’ is true iff it is not the case that if it had been that p it would not have been that q : in symbols, $p \diamondrightarrow q =_{df} \neg(p \Box\rightarrow \neg q)$. Defenders explicitly include Lewis (1973: 21), Bigelow and Pargetter (1990: 103), and Bennett (2003: 192), but Duality is implicitly assumed by most writers on ‘would’ counterfactuals (including Hawthorne in his presentation of the argument from inescapable clashes).⁴

¹This latter motivation from chances lies behind Hawthorne’s defence of (1), and behind Lewis’ presentation of a very similar worry: ‘if there would have been some minute probability of a quasi-miracle [a remarkable event], does it not follow that there might have been one?’ (Lewis, 1986a: 61).

²Of course, the version of counterfactual scepticism that Hawthorne is concerned with will remain live; but that is a difficulty for Lewis’ Humean similarity metric, not a general threat to counterfactual sentences.

³Though I use the symbols ‘ $\Box\rightarrow$ ’ and ‘ \diamondrightarrow ’, introduced by Lewis, I do not thereby intend to commit myself to his analyses—I use them merely as abbreviations.

⁴On Lewis’ account of the ‘would’ counterfactual, we get the derived truth condition that $p \diamondrightarrow q$ is true just when q is true at some close p -world. On that reading, Heller’s (1995) account (that $p \diamondrightarrow q$ is true just when q is true at some *close enough* p -world) is clearly closely related to Lewis’

Epistemic The ‘might’ counterfactual ‘if it had been that p , it might have been that q ’ is true iff it (epistemically) might have been the case that if it had been that p , it would have been that q : in symbols, $p \diamondrightarrow q =_{df} \diamond(p \Boxrightarrow q)$, where \diamond is an epistemic possibility operator of some kind. Defenders include Stalnaker (1981) and DeRose (1999).

Ontic The ‘might’ counterfactual ‘if it had been that p , it might have been that q ’ is true iff if it had been that p , it would have been that it (objectively) might have been that q : in symbols, $p \diamondrightarrow q =_{df} p \Boxrightarrow \diamond q$, where \diamond is an ontic possibility operator of some kind (e.g. logical or physical or nomological). Lewis (1986a: 63–5) defends this position.⁵

3 The Ambiguity Thesis

Having distinguished these theories, we must also decide whether we apply one of these semantic theories across the board, or whether some ‘might’ counterfactuals are governed by one theory, some by another. If the latter, ‘might’ counterfactuals are *ambiguous*. The most famous defender of ambiguity is Lewis (1986a: 64), where he argues that many ‘might’ counterfactuals are ambiguous between a reading in accordance with Duality, and a reading (which he calls the ‘would-be-possible’ reading) that is a version of the Ontic thesis. But one could conceivably regard ‘might’ counterfactuals as ambiguous between any of our readings, not just those two.

What can be said in favour of the ambiguity thesis? Lewis himself uses it to defuse the argument from inescapable clashes, by claiming that the argument equivocates. Certainly it is obvious that (3) is false according to the duality thesis: explicitly substituting the claim of Duality we get

$$(4) \quad (p \Boxrightarrow q) \wedge \neg(p \Boxrightarrow \neg\neg q),$$

which is obviously false. But, Lewis says, the second step of the argument is no longer motivated. In particular, the argument from ‘if it had been that p , there would be some chance that q ’ to ‘if it had been that p , it might have been that q ’ is

Duality thesis.

⁵I think it is more widespread than this single citation suggests. Bennett (2003) defends a view that is closer to Ontic than he perhaps admits, and (in personal communication) Daniel Nolan and Robbie Williams have both suggested a similar view to me. The view should be tempting to anyone who regards the meaning of the ‘might’ counterfactual as compositionally related to the meaning of ‘might’ and the semantics for ‘would’ counterfactuals.

valid on an ontic reading of the ‘might’ counterfactual—this is particularly obvious if, as many have suggested, ‘there is some chance that q ’ is a kind of (graded) ontic possibility operator itself (Mellor, 2000). So, Lewis thinks, one reading of (3)—namely, (4)—is false, but another reading (5) is true:

$$(5) \quad (p \Box \rightarrow q) \wedge (p \Box \rightarrow \Diamond \neg q).$$

So ‘we dare not treat ‘‘there would be some chance of it’’ and ‘‘it would not happen’’ in general as incompatible’ (Lewis, 1986a: 65).

While Lewis’ deployment of the ambiguity thesis serves to defuse the argument from inescapable clashes, it carries with it significant costs, and these costs look like they will apply to the use of any potential ambiguity thesis. To begin, Lewis’ use of it looks a little ad hoc, with the charge of equivocation wheeled out just to resolve a puzzle with his invocation of quasi-miracles, and no systematic evidence given for the two radically different meanings of ‘might’ counterfactuals. Relatedly, Lewis gives no guide as to when to deploy the ontic reading of the counterfactual, and when to use the duality reading: when counterexample threatens? When chances are explicitly mentioned? This will be a general difficulty for response to counterfactual scepticism that claims an equivocation is occurring. In fact, to move to our final worry, there seems to be considerable evidence that there is no equivocation here. If there were, (5) would be a perfectly acceptable reading of (3), and if we could contextually emphasise the right factors, we should be able to get ordinary speakers to assent to (3) in its ontic reading. Yet, as DeRose (1999: §13) points out, ordinary speakers are extremely reticent to assent to (3) in any context. Insofar as the ambiguity thesis gives an acceptable reading of (3), it makes the clash in (3) escapable, which is already a strike against the semantic adequacy of the proposal (and at least a preliminary strike against the ontic account involved in the ambiguity—of which more in §12).

None of these objections are knock down: the first two are requests for more detail about how and when the ambiguity operates, the third a worry that could potentially be explained away by pragmatic factors. Even so, I suggest they place some pressure on us to give a non-ambiguous account of ‘might’ counterfactuals if one can be given. (I return more favourably to the topic of ambiguity in §14.)

4 Lewis’ Argument for Duality

Duality is an appealing thesis, not least because the natural and compelling duality between necessity and possibility ($\Box p =_{df} \neg \Diamond \neg p$) is mimicked by the duality

between the corresponding counterfactual claims, as we might hope.

In fact, Lewis has an argument that Duality is the only acceptable account of ‘might’ counterfactuals. Lewis (1973: 80–1) invites us to consider this claim:

- (6) If I had looked in my pocket, I might have found a penny.

The argument is simple: if I did not in fact look, and there was no penny to be found, then this counterfactual is intuitively false, which is the verdict that the Duality theory delivers (since in all the closest worlds where I do look, I don’t find a penny). But, Lewis claims, neither the Epistemic nor Ontic theories (nor a number of other, less attractive, views) can explain why this sentence is false. For while ‘if I had looked (look), I would have found a penny (penny)’ is false, it is still compatible with all I know, and so $\diamond(\text{look} \square \rightarrow \text{penny})$ is true. Similarly, if I had looked, penny would have been false, but not necessarily false; so $\text{look} \square \rightarrow \diamond \text{penny}$ is true. If this argument succeeds, Duality is the only attractive theory remaining.

Yet Lewis must eventually have found something wrong with this argument, because as we’ve seen he winds up thinking a version of the Ontic approach is viable. The ambiguity thesis is one possible explanation; it may be that the Ontic approach still fails for (6), but that doesn’t mean it fails across the board. But I think we can do better in explaining why the Ontic theory can explain the falsity of (6), along the following lines. Lewis’ original objection to the Ontic theory is fine if the possibility operator occurring in the consequent is logical or metaphysical, and Lewis claims that these objections ‘persist for other sorts of possibility also’ (81). Yet it is by no means clear that it is physically possible for a coin to appear there, given that nearby worlds should match the way the actual world was up until just before I didn’t look in my pocket. Nor is it clear that in all the nearby worlds in which I look, there is any objective chance of a coin being in my pocket, unless we allow for some gratuitous differences between actuality and the closest worlds. These observations do indicate that the kind of ontic theory we accept must involve a kind of physical or probabilistic possibility in the consequent, so that whether p is possible at w will depend quite intimately on other matters of fact at w . Intuitively this is perfectly acceptable.⁶

⁶Perhaps some uses of ‘might’ counterfactuals (those with particular stress on ‘might’, for example, ‘I guess if the cold war hadn’t ended, there *might* be permanent manned bases on Pluto—but I don’t think it likely’) involve a weaker possibility, but those certainly seem to be exceptional in uses of ‘might’ counterfactuals. Perhaps these exceptional cases might be better rendered just as $\diamond(p \wedge q)$: nothing about their conditional form seems essential to what they are supposed to communicate, which is just that p doesn’t logically exclude q .

What about the Epistemic thesis? The first point is that on the Epistemic reading, there are perfectly good uses of (6), and it is *not* plainly false after all (DeRose, 1994). Secondly, even if the strong intuition that *sometimes* (6) is false means that we admit some oddity concerning the good uses of (6) (perhaps it is assertible while not true), there remains an Epistemic response to Lewis' argument. Stalnaker says that the relevant reading of \diamond should be 'not what is, in fact, compatible with my knowledge, but what would be compatible with it if I knew all the relevant facts' (1981: 101). In a similar vein, DeRose suggests that $\diamond p$ should be understood as

(1) no member of the relevant community knows that p is false, and (2) there is no relevant way by which members of the relevant community can come to know that p is false. (1991: 593–4)

Whatever the unclarities of these two proposals for \diamond , it is clear that neither of them will make (6) come out true. On Stalnaker's account, it is clearly relevant to (6) that there is no coin in my pocket, whether or not I currently know that, so it is not possible in his sense that I would find one if I looked. Similarly, there is a relevant way of finding out whether there is a coin in my pocket; so according to DeRose there is no epistemic possibility of 'look $\square\rightarrow$ penny', and (6) is false. As in the Ontic case, these results place constraints on the form of an acceptable Epistemic account of 'might' counterfactuals, and a flat-footed account of \diamond as 'for all I know, ...' can't be correct. Yet the Epistemic response is still viable.

I conclude that Lewis' direct argument for Duality is far from conclusive. I go on to argue that, in fact, Duality does much worse than its rivals in accounting for some uses of 'might' counterfactuals.

5 Against Duality I: The Argument from Bias

Consider a biased coin. If H is the event of that coin coming up heads on a toss, let $\text{Pr}(H) = 0.8$. Given this piece of information about the coin, it seems quite likely that if the coin were tossed, it would come up heads. It also seems quite likely that if I were to bet on H , I would win. Both the preceding judgments of likelihood seem to me intuitively plausible, and don't seem to require any intuition pumps or other stage setting to command our assent.

So much for what *would* happen on a bet or a toss. If we look at what *might* happen, similar intuitive judgments can be generated. For it seems very likely that if the coin were tossed, it might come up tails (though it doesn't seem very likely that if it were tossed, it would come up tails). And it seems very likely that if I were

to bet on heads, I might lose (though it doesn't seem very likely that if I were to bet on heads, I would lose). Again, these homely judgements command our intuitive assent.⁷

Yet taken together these rather elementary observations pose a problem for Duality. I earlier maintained, in the case of the biased coin, that

(7) It is likely that if the coin were tossed, it would land heads.

So, given that logically equivalent propositions get the same probability, Duality must similarly endorse the proposition that

(8) It is likely not the case that, if the coin were tossed, it might land tails.

So, by the laws of probability (that a proposition is unlikely to the extent that its negation is likely, for $\Pr(\varphi) = 1 - \Pr(\neg\varphi)$),⁸

(9) It is unlikely that if the coin were tossed, it might land tails.

But the denial of (9) is precisely what was endorsed one paragraph ago. Exactly similar reasoning will end in us endorsing the claim that it is unlikely that if I were to bet on heads, I might lose, again contrary to the supposition above. Something has gone seriously wrong.

Note immediately that this argument does not threaten the Ontic and Epistemic theories about 'might' counterfactuals, because, on both of those theories, nothing problematic about 'might' counterfactuals follows from (7).⁹

6 Against Duality II: Understanding Likelihoods of Natural Language Counterfactuals

Perhaps I have gone wrong in interpreting the natural language sentences. Perhaps the correct form of 'it is likely that if the coin were tossed, it would have landed heads' is not (7), but rather

(10) If the coin were tossed, it would likely land heads.

⁷Though some speakers I've consulted report some initial discomfort with the combination of 'likely' and 'might'. To my ear these sentences sound fine—it is true that if the coin were tossed, it might land tails (even despite its bias), and the involvement of 'might' makes this a very weak claim, and thus likely. (This is related to the observation that claims of mere possibility of φ are highly likely regardless of how likely φ itself is.)

⁸One might think that 'likely' isn't binary in this way; even in that case, the argument goes through equally well with 'more likely than not', or 'greater than 50% probability'.

⁹In fact, on an epistemic reading of 'likely', (7) just is the 'might' counterfactual 'if the coin were tossed, it might land heads', which is clearly compatible with our earlier remarks.

If (10) is the correct reading, the Lewis account allows us to derive, not (8), but

(11) It is not the case that, if the coin were tossed, it might likely land tails.

And (9) does not follow from (11), dissolving the apparent problem.

Yet there are several problems with this proposal. The first is that the reading in (10) seems quite remarkably artificial. The natural language sentence appears to attribute a likelihood to a counterfactual conditional proposition. Almost everyone who works on counterfactual conditionals thinks that they do express propositions, including Lewis (1973: 46–7).¹⁰ If they express propositions, they are the kinds of things which can have probabilities. Since the natural language sentences look to attribute probabilities to the counterfactual conditionals, and counterfactual conditionals are the kinds of things that can legitimately have probabilities, there seems no reason to perversely go against the most obvious and straightforward reading of the natural language sentences as in fact attributing probabilities to the counterfactual conditionals (certainly not as our first response to an apparent puzzle for a merely philosophical theory). What better way could there be for us to express our confidence in the conditionals than in the form of words used above and analysed as in (7)? And our utterance of conditional claims, and the fact that our deliberate action betrays our commitment to those conditionals as used in hypothetical reasoning, seem together to suggest that we are in many cases directly confident in the conditionals, not merely conditionally confident in their consequents.¹¹

If one decides to accept the deviant analysis of (10), other problems emerge. Let our biased coin have been tossed at t , and let H_t be the proposition that the coin landed heads at t . Then, presumably, prior to t , $\Pr(H_t) = \Pr(H) = 0.8$. But what should we say of the probability of H_t after t ? In this case, two kinds of probability emerge, and diverge from one another: the objective chance of H_t goes to 1 or 0, while our subjective credence might, if we didn't yet know the outcome, remain at 0.8. Now consider the situation when I bet after t , remaining in ignorance of the outcome; it still seems intuitive to claim that it is likely that if I were to bet on heads, I would win. On the deviant reading, that must be read as

(12) If I were to bet on heads, it would be likely that I win.

¹⁰Edgington (1995) may be an exception.

¹¹Incidentally, sentences of the form 'If — then it is likely that —', like (10), are a couple of orders of magnitude more common (in Google hits) than sentences of the form 'It is likely that if — then —', like (7). This is evidence that (10) is permissible; but it strikes me as no evidence that (10) is the 'real' reading, since ordinary speakers do not take (7) to be ungrammatical; that (10) and its ilk are common is no grounds to take the surface form of (7) to be drastically misleading as to its logical form.

But (12) is a very puzzling claim. If ‘likely’ in the consequent is an objective probability, then in the situation when the coin actually lands tails, as it might, (12) is false: for the objective chance of heads is 0. Yet the claim ‘it is likely that if I were to bet on heads, I would win’ can still be true in that situation, which means that the objective reading of ‘likely’ cannot be correct, and it must be read as a credence. But when read as a credence, (12) still says something puzzling: that if I were to bet on heads, then my credence in winning would be high. But why should my credence in winning counterfactually depend on whether or not I bet (except in the trivial sense in which I should never have bet had I not been confident of winning)? In general, it seems perfectly possible that we can be quite confident in ‘ $\varphi \square\rightarrow \psi$ ’ without thinking in any way that, were φ true, that would directly alter our credence in ψ . Of course were we to *find out* that φ was true, our typical reaction given our confidence in the conditional should be to increase our confidence in ψ ; but the claim here is that our credence directly counterfactually depends on the truth of the antecedent, which is rarely true. So I submit that neither the objective or subjective reading of ‘likely’ in the consequent gives the right reading in our case, and hence that the deviant reading (10) cannot be correct.

Some final evidence that (10) cannot be the correct reading is that, if we are serious about Duality, we should think that the ‘might’ and ‘would’ counterfactuals behave similarly with respect to likelihoods, so that the correct reading of ‘it is likely that if the coin were tossed, it might land tails’ should be

(13) If the coin were tossed, it might likely land tails.

And (13) directly contradicts (11). So, whether we read ‘it is likely that if the coin were tossed, it would land heads’ as (7) or (10), we end up in a contradiction if we treat the corresponding ‘might’ counterfactual as Duality requires.

This problem of embedded counterfactuals in likelihood contexts is quite general. People may not be inclined to assent to

(14) If the coin were tossed, it would land heads,

just because they think it likely. Yet they are inclined to deliberate and act (for example, in deciding how to bet) *as if* they regarded (14) as true just because they think it likely. In light of this observation, it seems quite clear that in many contexts in which counterfactuals are central, in hypothetical reasoning most prominently, it is really the fact that those counterfactuals are thought likely that explains why people act on them and even assent to them. I take it that belief is a matter of degree,

modelled by probabilistic credences; hence the failure of Duality to play nicely with probabilistic operators would undermine most uses of ‘might’ counterfactuals.

Taking this likelihood-embedding problem together with the fact that Duality cannot (without the aid of Ambiguity) avoid the argument for counterfactual scepticism from inescapable clashes, I conclude that Duality cannot be the correct account of the ‘might’ counterfactual.

7 The Epistemic Theory

The epistemic theory starts from the observation that ‘might’ is typically used, outside of conditional contexts, to express epistemic possibility. An utterance of ‘John might have come to the party’ certainly communicates that the speaker thinks it possible that John was at the party, but since it is not felicitous in contexts where the speaker knows whether or not John was at the party, that possibility must be epistemic.¹² At this stage we can contrast typical uses of ‘could’: ‘John could have come to party’ also expresses a kind of possibility, but it is assertible when the speaker knows that John in fact did not.

There is some dispute over just how to understand epistemic possibility; of the two similar options canvassed earlier in §4 (page 7), I will focus my attention on the account of DeRose (1991). It is worth noting that other theories of epistemic modals, such as recently popular relativist proposals (Egan *et al.*, 2005), might give rise to quite different discussions at this point.

With this claim about ‘might’ in place, the epistemic theory emerges when one insists that ‘might’ in conditionals plays precisely the same role as ‘might’ outside of conditionals (Stalnaker, 1981: 99). So ‘ $p \diamondrightarrow q$ ’ is not some kind of indivisible semantic unit, but is rather to be explained jointly in terms of the semantics of conditionals and in terms of the semantics of ‘might’. If so, then most uses of ‘might’ counterfactuals will be used to express the speaker’s recognition that the corresponding ‘would’ counterfactual is epistemically possible.

The preceding argument obviously rests on the claim that ‘if p then it might be that q ’ is synonymous with ‘it might be that if p then it would be that q ’ (so the appearance of ‘might’ in the consequent of the former is misleading). This claim can be supported when one considers the sentence

¹²Yet ‘John might have come to the party’ is fine as a past tense claim when the speaker now knows whether John did; the explanation for this phenomenon will have to wait until §10.

(15) Had this plane just been diverted to Montreal, I might still be in New York in four hours.

(15) is intuitively false when uttered by me on a plane in the mid-Atlantic; but because nothing I know at that point rules out my being in New York in four hours, the conditional ‘if this plane was just diverted, then \diamond I’m in New York in eight hours’ is true—so the reading with ‘might’ in the consequent gives the wrong truth conditions on ‘might’ sentences. (Exactly similar phenomena exist for ‘must’ conditionals.)

The first major test of the Epistemic theory is whether it can account for the firm intuition that sentences of the form (3) are false: that is, can the Epistemic theory account for the phenomenon of inescapable clashes? At first glance, it might seem that it cannot. If we explicitly substitute the most basic account of epistemic ‘might’ into an instance of (3) (an account similar to Stalnaker’s, and entailed by DeRose’s more complicated proposal (p. 7)), we could get the following:

(16) If it had been that p , it would have been that q ; and yet for all I (and my mates) know (or could plausibly find out), if it had been that p it would have been that $\neg q$.

I agree that (16) is odd, but its oddness doesn’t seem to preclude its truth. It may well be that the first conjunct is true, and the asserter of (16) gets that right (perhaps by a lucky guess), while fully recognising that they don’t know that the first conjunct is right and thus asserting the second conjunct. There is some tension between the conjuncts of (16), but it is pragmatic, not semantic tension (as in Moore’s paradox). So (16) is true, so instances of (3) can be true; yet ordinary speakers’ thoroughgoing rejection of (3) shows that (16) can’t be the right reading and hence that the Epistemic theory is incorrect.

This is a compelling argument against the Epistemic account only if the phenomenon of inescapable clashes required that those clashes be semantically explicable. But no such requirement exists; the phenomenon to be explained is that ordinary speakers reject instances (3), not that they definitely think it false and defective for semantic reasons. It could well be that the reason they reject (3) is because an utterance of (3) invariably communicates (implicates) a false proposition without itself being invariably false (this is certainly what seems to be going on in Moore’s paradox). A sketch of how this implicature works might be this: a genuine assertion of (3) asserts not just the content of the sentence, but also that the speaker takes themselves to know the content of each conjunct. But the claim

that the speaker knows the first conjunct of (3) is not consistent with the claim that, for all the speaker knows, the first conjunct of (3) is false.¹³

If an assertion of (3), and therefore of (16) on this picture, is pragmatically defective because it implicates a falsehood, that explains the *phenomenon* of inescapable clashes while avoiding the *argument* from inescapable clashes to the counterfactual sceptic's conclusion: for if (3) is true (though unassertable), we needn't reject either conjunct, and in particular needn't reject the first conjunct, the ordinary counterfactual claim. So there isn't a good argument for 'might' counterfactuals to counterfactual scepticism, on the epistemic view.

We are still faced with some unpleasant conclusions: if we take ourselves to know the second conjunct of (3), we cannot take ourselves to know the first conjunct of (3) (even though, on the Epistemic reading, that conjunct can be true). This leads to what we might call *weak counterfactual scepticism*: the thesis that, even if they are true, ordinary 'would' counterfactual claims can't be known if the corresponding 'might' counterfactuals are known. For the Epistemic theory, knowing the second conjunct of (3) does involve knowing the nature of one's own epistemic state, and it could be argued that this is rather more difficult to come to know than might have been expected. Yet even if the second conjunct of claims like (3) are difficult to come to know, that doesn't really remove the worry that whenever we do come to know them, the corresponding 'would' counterfactuals can't be known, and arguably can't be believed or acted on either.¹⁴ Weak counterfactual scepticism threatens the Epistemic account precisely because in (3) the epistemic 'might' is an operator on a conditional sentence that is incompatible (when the antecedent is non-contradictory) with the original 'would' counterfactual, so that if the latter is taken to be known, the former is known false. Of course, it is this very feature that makes the clashes seem so inescapable, and to abandon it appears to abandon any attempt to explain the fundamental phenomena of clashes. In any case, the original argument for counterfactual scepticism is invalid on the Epistemic view, and weak counterfactual scepticism is perhaps a less compelling or threatening position.

¹³As standard logics for counterfactuals validate $p \Box \rightarrow q \vdash \neg(p \Box \rightarrow \neg q)$ when p is not an impossibility.

¹⁴It may be, though I cannot develop the thought here, that a contextualist semantics for 'knows', or something similar, can dissolve this problem: in many ordinary circumstances the conditions on knowing the first conjunct of (3) are satisfied, but those circumstances in which the second 'might' conjunct of (3) is made salient prevent the first 'would' conjunct from being known.

8 Objective Uses of ‘Might’

A serious threat to the epistemic position, however, is the fact that many uses of ‘might’ seem not to involve epistemic states at all. Stalnaker explicitly acknowledges this, and also allows for a non-epistemic use of ‘might’ (Stalnaker, 1984: 143–6). This proposal is more appropriately treated as an Ontic approach combined with adherence to the Ambiguity thesis, and DeRose’s strong arguments against this package will be treated in §12. DeRose however makes no allowance for this kind of Ontic theory, and undertakes to explain away these apparently non-epistemic uses of ‘might’ counterfactuals. At this point the full resources of his sophisticated account of epistemic ‘might’s are brought into play.

The kinds of cases that are of concern for a purely epistemic theory are these:¹⁵ imagine two fully informed quantum physicists, who are considering performing an experiment on an electron with two possible outcomes, ‘spin up’ and ‘spin down’, yet such that the prior state of the electron doesn’t determine the outcome—it is a genuinely indeterministic situation. Funding constraints prevent them from running their experiment; they may regretfully say:

- (17) If we had performed our experiment, we might have got the result ‘spin up’.

According to the epistemic account, (17) is to be analysed as something like

- (18) For all we know and can relevantly find out, if we had performed our experiment, we would have got the result ‘spin up’.

The immediate objection to this analysis is that (18) is arguably false. For if the situation is genuinely indeterministic, then there doesn’t seem to be any fact of the matter about what *would* happen if the experiment were run—and on that basis we should conclude that any ‘would’ counterfactual which claims a determinate outcome would occur if the experiment were run is indeed false—that’s what indeterminism is all about.¹⁶ Our scientists it seems are thus in a position to know that the would-counterfactual corresponding to (17) is not true, and so (17) cannot be compatible with their knowledge, as (18) requires.

It is possible to undermine the preceding argument—one could, for instance, point out that it is uncomfortably similar to the suspicious argument for counterfactual scepticism in §1. Yet even if we did so, and became convinced that (18) is true

¹⁵A similar case appears in DeRose (1999: §9).

¹⁶Jeffrey (1992: 193) makes a similar point.

when (17) is, (18) seems not to be a good account of *why* (17) is true. It is perhaps more natural to think that the objective facts of the situation make both (17) and (18) true, the latter in a derivative fashion because it is not possible to know what isn't objectively settled. (18) only comes out true because we are guaranteed by the physical indeterminism that there is no way at all for our physicists to come to know the consequent of the embedded counterfactual.

This point can be brought out more clearly by considering a closely related case: consider

- (19) If an electron had been in situation *E*, then it might have exhibited 'spin up' behaviour.

If 'situation *E*' is the experimental situation in (17), then it is natural to think that (19) and (17) are both true for the same reasons, namely the indeterministic nature of situation *E*. Yet this indeterminism does not entail the epistemic reading of (19), since we can consider a world without any epistemic agents but otherwise sharing our physics, in which electrons still behave similarly in situation *E*, yet no one exists to know (or fail to know) anything. (19) could be true in a world which had no epistemic agents; yet no claim about what is compatible with the knowledge of epistemic agents could be (non-trivially) true in such a world. Again this is not decisive, for we may claim that (19) is really true only *of* (not in) a world which contains no one to entertain it, and an epistemic claim can be true of a world without agents (from our perspective, for example).

Though not decisive, these observations do make problems for the epistemic theory which simply do not arise for the ontic theory. It is easy to see that if (17) means (20) there is no problem posed by objective indeterminism:

- (20) If we had performed our experiment, we would have had some chance of getting the result 'spin up'.

9 'Could' Counterfactuals

More decisive problems for the epistemic theory arise from a different direction, drawing on considerations related to the neglected topic of 'could' counterfactuals. I argue firstly, in the present section, that there is a distinctive non-epistemic reading of 'could' in such counterfactuals, when understood correctly; and I argue secondly (in the following section) that sometimes we use a 'might' counterfactual to express one of these non-epistemic 'could' counterfactuals. From these two

claims it quickly follows that there are non-epistemic uses of ‘might’ counterfactuals, contrary to the epistemic theory.

The first claim depends on there being an account of ‘could’ counterfactuals, sentences such as:

(21) If he had shot from point blank range, Assassin could have killed Mr Smith.

Yet such counterfactuals have received almost no attention in the literature;¹⁷ and it is not at all clear that what we say about the semantics of ‘might’ counterfactuals can explain the behaviour of ‘could’ counterfactuals. In what follows, I present an account of ‘could’ counterfactuals that I think is adequate to the phenomena. The obvious fact that ‘could’ is the indicative past of ‘can’ makes it convenient to take the analysis of ‘could’ in two parts, first looking at the role of ‘can’, and then at the contribution of the counterfactual construction.

Unlike ‘could’, ‘can’ has received considerable attention in the literature. The standard account is something like this: a sentence of the form ‘can p ’ is true just in case p is compatible with some contextually salient background considerations b (Kratzer, 1977; Lewis, 1983; 1986b). These background considerations can be left for context to supply, but may be made explicit by means of a phrase like ‘in view of’, as in (22):

(22) In view of his wealth, Bill Gates can afford a new boat.

But just reading ‘can p ’ as ‘ p is compatible with background conditions’ is not the whole story, for as linguists have emphasised, ‘can’ is very commonly used in the so-called ‘dynamic’ sense, that is, to attribute an *ability* (Palmer, 2001: 79–80). And this fact seems to make at least some uses of ‘can in view of’ sound odd to say the least:

(23) ?? In view of what I believe, James can polevault 6m.

The explanation for (23) is that, while it is true that polevaulting 6m is compatible with my beliefs, it doesn’t make it true that the person involved has the ability to do so, as I may be very ignorant. What matters, it seems, is what the object in question is like intrinsically, and it is the compatibility of p with the intrinsic nature of the subject of p that makes for the truth of an ability claim. There are, of course, further facts that can be contextually salient and while are themselves incompatible with p , which can thus undermine the truth of an ability claim even under circumstances

¹⁷Austin (1970) is an exception, though even here it is expressly concerned mostly with the role of ‘could’ conditionals in attributions of free action.

in which the object in question has a nature that is intrinsically compatible with p . Consider here Lewis' examples (Lewis, 1986b: 77):

- (24) a. In view of his larynx and nervous system, David Lewis can speak Finnish.
b. In view of his lack of training, David Lewis cannot speak Finnish.

The important thing to note about (24b) is that, intuitively, the facts about Lewis' larynx and nervous system remain relevant even in the context where the lack of training is also relevant; my basic claim is that to make correct sense of ability ascriptions, the shifting contextually relevant background facts must always minimally include some facts about intrinsic natures of the object to which the ability is ascribed. This proposal is compatible with all the evidence that Lewis and Kratzer use to support the context-sensitivity of 'can' while ruling out as irrelevant (at least for the dynamic use of 'can') some other possible background relevant facts, say those about my belief.

The upshot of this discussion is that 'can $P(o)$ ' means something like 'the way that o intrinsically is, and the relevant facts, are together compatible with o 's being P '. Given that ability ascriptions are relatively resistant to small shifts of context, there is apparently another reading, on which 'can $P(o)$ ' means something like, '*ceteris paribus*, P is compatible with o 's intrinsic properties'. These two proposals amount to much the same thing in practice, as if the relevant facts are enough to rule out o 's being P , then in this case *ceteris* will not be *paribus*.

Given that understanding of 'can', I now turn to the interaction with the conditional construction. One intuitively valid inference is from 'I can φ ' to 'If I were to try, I could φ '. In this case, the antecedent is in some sense irrelevant, since the ability to φ is not explicitly conditional on circumstances. But oftentimes we regard our abilities as requiring special circumstances to obtain in order that we might exercise them, and it is in specifying under which conditions our abilities can manifest that 'could' counterfactuals have their most significant use. So the 'could' counterfactual emerges naturally as a kind of conditional ability ascription: what our abilities would be if the circumstances were appropriate. The natural use for such a construction is in those contexts where q is ruled out by known facts, and yet under conditions of q , an ability currently lacked might well be manifested. To continue our earlier example:

- (25) If David Lewis had been given proper training, he could have spoken Finnish.

Taking this observation seriously, we arrive at the following proposal:

Ability The ‘could’ counterfactual ‘if it had been that p , it could have been that q ’ is true iff under the counterfactual assumption that p , ‘can q ’ would have been true: in symbols, $p \Box \rightarrow \blacklozenge q$, where ‘ \blacklozenge ’ represents the ‘can’ sentential operator. (Note that the context-sensitivity of the counterfactual combines straightforwardly with the context-sensitivity of ‘can’.)¹⁸

This proposal makes sense of (21): Assassin has no unconditional ability to kill Mr Smith, but certainly in the circumstances where he shoots from point blank range, he possesses that ability in virtue of his natural aptitude for assassination.

Other proposals for understanding ‘could’ counterfactuals do not succeed as well. The ‘can’ operator cannot function as epistemic ‘might’ does, with the \blacklozenge taking scope over the whole conditional (p. 11). For

(26) If the coin were tossed, it could land heads

is true, while the corresponding ‘would’ counterfactual is false; hence

(27) \blacklozenge if the coin were tossed, it would land heads.

is also false, because the constituent claim is not compatible with the relevant background of actual fact about coin tossing.¹⁹ Obviously, if the argument I make below about the use of ‘might’ to mean ‘could’ goes through, the Ability theory of ‘could’ counterfactuals already is in some tension with the Epistemic theory of ‘might’ counterfactuals, as even if \blacklozenge is an epistemic operator it only takes scope over the consequent, not the whole conditional as the Epistemic theory demands.

But this is a moot point, for it is quite clear that \blacklozenge is not an epistemic operator. It is a relative modality, which takes a set of contextually relevant facts, some of which must include facts about the intrinsic nature of subject of the proposition in the scope of the operator. Epistemic readings of these relevant facts cannot do justice to the fact that, even in cases where our state of knowledge about the relevant facts is compatible with $P(o)$, ‘can $P(o)$ ’ may be false because o is in fact such as to not be possibly P in some ontic sense of possibility (arguably physical possibility). For instance, for all Lois Lane knows, it isn’t the case that Clark Kent

¹⁸So ‘could’ here is taken to be functioning as ‘it could have been that’, rather than modifying the verb of q —see Fara (unpublished: esp. footnote 4).

¹⁹To put it another way, since the important fact about coin tosses and other gambling systems is that there is no determinate fact of the matter about how they will come up if trialled, both the counterfactuals ‘Toss $\Box \rightarrow$ Heads’ and ‘Toss $\Box \rightarrow$ Tails’ must come up false, and hence *cannot* come out true if the gambling system is to deserve the name. So (27) is false.

can fly; nevertheless she is wrong about that, because Clark Kent is in fact such as to be able to fly.

The Ability analysis governs only ‘could’ counterfactuals. The existence of sentences like (28) is quite clearly an evidential or epistemic use of ‘could’, not the dynamic ability sense that I’ve defended:

(28) That could be Elyse; she’s supposed to arrive about now.

This does seem to indicate that ‘could’ is not univocal in counterfactual and non-counterfactual contexts; all my argument needs is that many could-counterfactuals are clearly non-epistemic, which I think is evident.

10 ‘Might’ and ‘Could’ Counterfactuals

Various sorts of evidence are available for the claim that sometimes sentences of the form ‘if it were that p , it might be that q ’ are used to express the conditional proposition that if it were that p , it could be that q . The most natural starting point is that there are considerable intuitive similarities between ‘might’ and ‘could’ because both fundamentally express possibility. Since both words have this relation to ‘possibly’, and ‘possibly’ is systematically ambiguous (between epistemic and ontic possibility, for example), it would not be surprising if all of the possibility-expressing modal auxiliaries could be used similarly and without clear distinction. That ‘might’ and ‘could’ are sometimes used to express the same content is intuitively obvious; but it is worth seeing what arguments can be given.

Indirect evidence for this thesis is that many competent speakers of English regard *deontic* modal auxiliaries as interchangeable (Palmer, 2001: 80):

- (29) a. You could try complaining to the ombudsman.
b. You might try complaining to the ombudsman.

Insofar as these examples demonstrate that ‘might’ and ‘could’ (or ‘can’) may be used to express the same claim, it is evidence for our thesis. Yet both sentences in (29) express a deontic sense, while, as I discussed above, the canonical sense for ‘could’/‘can’ is to express an ability. It turns out that this ontic or dynamic reading of might is also possible: so Stalnaker argues that

might sometimes expresses some kind of non-epistemic possibility. *John might have come to the party* could be used to say that it was within John’s power to come, or that it was not inevitable that he not come. (Stalnaker, 1981: 99)

As he continues, Stalnaker in fact moves directly from discussing a ‘might’ claim to discussing a ‘could’ claim in such a way as to suggest that he regards them as in some sense synonymous; that this move is natural and unobjectionable in the course of his argument is evidence that ordinary speakers can use either formulation to express the ability claims in question. Stalnaker’s further claim that there is no special conditional use of ‘might’ then leads to the conclusion (with the results of the last section) that we could use either ‘might’ or ‘could’ interchangeably in some conditional contexts. For instance, these seem to express the same claim, namely that if she hurried, Jeff had it in his power to arrive at school by the time of utterance:

- (30) a. If he were hurrying, Jeff might be at school already.
b. If he were hurrying, Jeff could be at school already.

Particularly interesting examples appear when we explicitly emphasise the relevant abilities and derive the ‘could’ and ‘might’ claims from them:

- (31) a. Commodity prices can fall quickly; if they were to fall too quickly, then they could have an impact on the whole market.
b. Commodity prices can fall quickly; if they were to fall too quickly, then they might have an impact on the whole market.

It seems implausible that the examples in (31) really express different propositions; rather, what might happen in that situation is precisely what could happen. If we are aware that it might have an impact, then we are in a position to infer that it could have an impact; similarly, if we are aware that it could have an impact, then we can infer that it might.²⁰

One final piece of evidence is that there is a phenomenon of clashes for ‘might’ and ‘could’ counterfactuals:

- (32) If it were that case that p , it might be the case that q ; yet if it were that case that p , it couldn’t be that q .

So, for instance, many instances of (32) sound bad: for example, (33a). If ‘might’ could never mean ‘could’, we have no explanation for that.

²⁰Consider also the following conversation:

A: If the traffic was light, Spencer and Marnie could be home by now.

B: You’re right: they might be.

That this fragment is natural seems to indicate that no obvious distinction between the role of ‘might’ and ‘could’ appears in natural language.

- (33) a. If we'd left the gate open, the dog might have got out; yet if we'd left the gate open, the dog couldn't have got out.
 b. If we'd left the gate open, the dog could have got out; yet if we'd left the gate open, it isn't the case that the dog might have got out.

It is interesting that the opposite kind of clash to (32), as exemplified in (33b), is not inescapable: there are good uses of that latter sentence (a very obedient dog, for example).²¹ In any case, all that matters is that (33a) is a clash, for that shows that the 'might' conditional isn't plausibly true without the 'could' conditional being true; which indicates that the 'might' conditional implies the 'could' conditional and thus has an ontic use. Of course this is not knock-down—we wish to undermine a similar argument that derives from (3) the claim that 'might' and 'not-would-not' are synonymous—but that the clash be explicable is a basic desiderata on any theory, one which the current proposal easily meets.

Of course, various prescriptivists about language say that the use of 'might' to mean 'could' is incorrect: Austin calls it a 'vulgarism' (1970: 207). Such prescriptive judgements aren't supported by very much evidence, though the behaviour of 'could' and 'might' under negation is worth noting here, because of non-synonymous pairs like

- (34) a. Even if she trained, she could not run faster than him.
 b. Even if she trained, she might not run faster than him.

But that fact that 'might not' typically expresses 'possibly not', and 'could not' typically expresses 'not possibly', is far from decisive evidence for the claim that 'could' and 'might' *never* can be used in the same sense.

11 The Ability Theory of 'Might' Counterfactuals

The results of the preceding three sections suggest a distinctive kind of ontic theory: that 'might' counterfactuals are 'could' counterfactuals, and that the 'ontic possibility operator' I introduced when I first defined Ontic theories in §2 turns out to be the 'can' operator. Call this version of the ontic theory the *ability* theory of 'might' counterfactuals. Intuitively, the proposal has some plausibility even before we begin to look at cases: certainly those things which are able to happen might happen, in some straightforward sense of 'might', and the ability theory gives the

²¹This seems to be further evidence in favour of the claim that 'could', when used explicitly, can't be synonymous with 'might': otherwise there would be inescapable clashes and there aren't.

most straightforward explanation of this connection between ‘can’ and ‘might’—far more straightforward than any epistemic theory could hope to provide.

This thesis makes good sense of the ontic uses of ‘might’ that I considered in §8. For instance, (19) is rendered as

- (35) If an electron had been in situation E , it would have had the ability to exhibit ‘spin up’ behaviour.

(35) is straightforwardly true in virtue of the way that electrons are intrinsically, and seems to capture exactly the claim in (19). The ability of the electron to do exhibit either spin up or spin down behaviour allows us to think it might do either; if we thought otherwise we’d presumably reject the ability claim.

Interestingly, (17) raises some complexities, as straightforward reading of (17) would seem to ascribe the ability to the experimenters, not to the entities involved in the experiment! In this case the ability of the experimenters to (veridically) observe a certain result is obviously parasitic upon the more fundamental ability of the experimental apparatus to produce that result. And arguably if the experimenters had lacked certain visual abilities they wouldn’t have been able to see their result, so their abilities do seem to be relevant to some extent. In either case, the appropriate ability claims are ontic, not epistemic, and the appropriate reading seems to be something like:

- (36) If we had performed our experiment, we would have been able to observe the result ‘spin up’ (in virtue of the ability of the experimental apparatus to produce that outcome).

There seems to be an obvious sense in which the outcome of the experiment might have been either spin up or spin down because the outcome of the experiment was able to be either spin up or spin down. In general, the awkwardness of (37) seems to show a close connection between ‘might’ and ‘could’:

- (37) ?? If it had been that p , it might have been that q ; but it’s false that if it had been that p it could have been that q .

One issue from §8 is outstanding, which is that I already gave an ontic reading of (17) as (20), a claim explicitly involving chances, and it would be nice to show that (20) and (36) are compatible. This is actually fairly straightforward. Recall the discussion in §3 (p. 5), where it was suggested that chance ascriptions were a kind of graded ontic possibility. One may now ask: what kind of ontic possibility? Metaphysical possibility seems too weak, as we do not wish to assign a non-zero

chance to an event just because it is metaphysically possible. It seems that chances should have something to do with the best physics, but even physical possibility is too weak, as it is clearly physically possible that the past might have been different but that has no present chance. I claim that the appropriate kind of ontic possibility is the ability modality ‘can’, because it seems appropriately stronger than physical possibility yet does not reduce chance to merely actual phenomena.

In order for the present analysis in (36) to suffice, all I need to show is that the analysis in (20) entails (36). If q entails r , it is a theorem of counterfactual logic that $p \Box \rightarrow q \models p \Box \rightarrow r$, so for (20) to entail (36) all I need to show is that ‘there is some chance of $P(o)$ ’ entails ‘can $P(o)$ ’. This seems to be the case: a counterexample would be a situation in which o has some non-zero chance of satisfying P and yet is not able to, i.e. a situation in which P is not compatible with the other facts about o . But in that situation the chance will be conditional on those other facts, and the chance of P must therefore be zero, contrary to assumption. It is not possible therefore to have a non-zero chance of some event and yet that event be unable to be realised (?). I conclude that the present analysis of ontic uses of ‘might’ counterfactuals in terms of ‘could’ counterfactuals is compatible with, and indeed is entailed by, those accounts of ontic ‘might’ counterfactuals where the operator on the conditional is a probability, as in (20) but more importantly as in Lewis’s later ‘would-be-possible’ reading of the ‘might’ counterfactual (Lewis, 1986a: 64). Lewis agrees with the present claim that the counterfactual mentioning chances in the consequent entails the corresponding ‘might’ counterfactual.

Returning now to the problem of counterfactual scepticism, we can see that the Ability theory is able to avoid the sceptical conclusion. According to the Ability theory, (3) is to be read:

- (38) If it were that p , it would be that q ; and if it were that p , it would be that it can be that $\neg q$ (i.e. it could be that $\neg q$).

As ‘can’ in (38) functions like a possibility operator, in that $\blacklozenge q$ can be true even when q is false. As such, it can be true that $p \Box \rightarrow q$ while it is also true that $p \Box \rightarrow \blacklozenge \neg q$, which is all that is needed to make (38) true. Lewis makes much the same point about his ‘would-be-possible’ reading, and also observes that we cannot in general treat ‘it would not happen’ and ‘it can happen’ as incompatible without serious modification to the underlying similarity semantics for counterfactuals (Lewis, 1986a: 65). This observation derives from the elementary point that there can be unexercised abilities. Indeed, given our account of ability as deriv-

ing from intrinsic properties, there can be abilities that are prevented by external circumstances from ever being exercised (as Lewis also points out, 1986b: 77–8).

12 Inescapable Clashes and Ontic Theories

The main threat posed by the argument from inescapable clashes was that (3) seemed not only to be false, but also that its falsity apparently was demanded as a condition of adequacy on a semantics for ‘might’ counterfactuals. Any theory, like the Ability theory, on which (3) can be true is therefore obligated to give an account which explains why (3) sounds so horrible to ordinary speakers, bad enough that they invariably take it to be false even in cases in which it is true.

DeRose (1999: §13) thinks that ontic theories cannot explain this phenomenon (he focuses on Lewis’ ‘would-be-possible’ reading but I assume that it is supposed to apply, *ceteris paribus*, to other ontic theories). His argument is simple: if ‘might’ counterfactuals had the ontic reading, then there are contexts which the second conjunct of (3) can sound true. If we are in one of these contexts, we should be able to felicitously utter (3), because the second conjunct would take the true reading. But no instance of (3) can be uttered felicitously. So there doesn’t seem to be a context in which the ‘might’ counterfactual takes the ontic reading. In a sense, the Ontic reading makes it too easy for (3) to be the correct thing to say in certain circumstances.

DeRose’s argument also serves to make a preemptive strike against the Ambiguity thesis (which I will defend in §14). We might think that the ‘might’ conditional can sometimes take the ontic reading, and sometimes an epistemic reading, and it may just be that in the context of (3) the ontic reading *never* takes precedence, not even when we’ve made the ontic elements of the situation as contextually salient as we can. Yet an Ambiguity theory for some expression which says that one of the disambiguations fails to share one of the core features of the expression is *prima facie* inadequate. So this theory, where there is an ontic reading that can never be emphasised in precisely the situations we need, would be inadequate. It is incumbent on any defender of an ontic theory, whether regarded as part of an ambiguity theory or not, to explain (or explain away, hopefully) the phenomenon of inescapable clashes. In the next section, I will do just that for the Ability theory.

13 Inescapable Clashes and the Ability Theory

DeRose himself thinks that (3) can be true, yet any assertion of it pragmatically implicates a falsehood which explains why (3) sounds bad. So if the Ability theory can give a similar explanation, by DeRose's lights that should be enough to accept that the Ability theory can be an adequate semantics for 'might' counterfactuals.

I believe that there is a pragmatic explanation available; in fact, there are at least two. Before I give them, recall some facts about the standard semantics for the 'would' counterfactual: a counterfactual conditional is true just in case there is no contextually relevant situation in which the antecedent is true but the consequent false. For example, the original Lewis-Stalnaker proposal analyses contextual relevance in terms of global similarity between the actual situation and a counterfactual situation in which the antecedent obtains (Lewis, 1973; Stalnaker, 1968). Even if one disagrees with the details of the Lewis-Stalnaker proposal about similarity, it is certainly true that intrinsic facts about the objects involved in the antecedent propositions are prominent respects of similarity.

I argued earlier (p. 18) that the correct semantics for the 'could' counterfactual were given by the Ability analysis, so that 'if it had been that p , it could have been that q ' is to be understood as $p \Box \rightarrow \blacklozenge q$. This latter sentence is not equivalent to $\blacklozenge(p \wedge q)$ ('it can be that p and q '), for it may be the case that $\blacklozenge q$ is actually false, yet would be true under the counterfactual supposition that p . For instance, I can't actually dunk a basketball, yet if I had been a meter taller, I could do so easily. The difference between the two sentence emerges in this example because the antecedent p proposes a counterfactual variation in the intrinsic properties of the objects involved in p , and this variation supports a difference in the abilities of those objects: in our case, varying my intrinsic height supports a difference in my ability to dunk a basketball, even though actually, with my actual intrinsic properties, I lack this ability. This explanation of the difference generalises: the situations in which $\blacklozenge(p \wedge q)$ differs from $p \Box \rightarrow \blacklozenge q$ are just those where p involves a counterfactual difference in intrinsic properties (rather than extrinsic circumstances).

In other cases, I suggest, it is extremely easy to think that $\blacklozenge(p \wedge q)$ and $p \Box \rightarrow \blacklozenge q$ have the same truth value, and indeed to hear the 'could' counterfactual as expressing both claims. So, for instance, (39) rightly sounds very bad:

- (39) If there had been any oxygen, the match could have lit; but it can't be that there was both oxygen and a lit match (i.e. $(p \Box \rightarrow \blacklozenge q) \wedge \neg \blacklozenge(p \wedge q)$).

(39) says that the antecedent of the first conjunct is both compatible with the con-

sequent and incompatible: the match both has the intrinsic ability to light in extrinsically different circumstances and yet also lack the intrinsic ability to light in different circumstances. The badness of this sentence seems, on a cursory reading, to carry over to the superficially similar (40):

- (40) If I'd been quicker, I could have saved that extra run; but I can't have been quicker and saved the extra run.

On closer inspection, however, it's quite clear that (40) is fine: for example, one could use it to give an excuse for failing to save the extra run (i.e. even though someone faster than I am could save that run, I'm not actually that fast and I don't have the ability, in virtue of the way that I am, to be faster). But the distinction here is subtle, and it is quite easy for ordinary speakers to elide the difference and take (40) to be false and hence to think that the 'could' counterfactual in (40), like the one in (39), is to be reading as asserting the actual ability of p and q to be true, rather than a merely counterfactual ability.

The Ability theory says that a 'might' counterfactual $p \diamondrightarrow q$ can have the 'could' reading of $p \squarerightarrow \blacklozenge q$. Given the natural and understandable potential for confusion between the latter sentence and $\blacklozenge(p \wedge q)$, it is quite easy to think that a 'might' conditional $p \diamondrightarrow q$ expresses $\blacklozenge(p \wedge q)$. If this confusion happens, and it seems pretty clear that it can happen, (3) gets the following reading:

- (41) $p \squarerightarrow q \wedge \blacklozenge(p \wedge \neg q)$.

And (41) is typically false. Given the standard semantics for 'would' counterfactuals, the first conjunct of (41) is true iff all the situations in which things are similar to actuality insofar as they are compatible with p are situations in which q is true. For the second conjunct of (41) to be true, there must be some situations in which $\neg q$ is compatible with p , yet they cannot be relevant situations because in all relevant p -situations, q is true. But we do not regard ability claims as being made true by irrelevant situations; so in general the first conjunct can only be true if the second is false. The contrary claim is true also: the second conjunct can be true only if the first is false, for if the second conjunct is true, there are nearby situations in which p and $\neg q$.

Since (41) is false, and there is an understandable confusion whereby ordinary speakers can take (3) to mean (41), there is little wonder that (3) sounds terrible to ordinary speakers. This happens even though the semantics of (3) are in fact quite different to (41), and the Ability theory explains why (3) can be true. Call this the *Understandable Confusion* explanation for inescapable clashes.

There is a complementary explanation which is particularly apt in the case of (3). The assertion of the ‘would’ counterfactual that forms the first conjunct of (3) is tantamount to the assertion that any nearby situation in which p holds is also one in which q holds. Thereafter, any other ‘would’ counterfactual in which p is the antecedent will tend to assume this fact, and thus assume q as known background knowledge in p -situations. If q is a fact to be held fixed in the evaluation of counterfactuals with p as antecedent, then q will be held fixed in the evaluation of the ‘could’ counterfactual $p \Box \rightarrow \blacklozenge \neg q$. Understandably, if q is held fixed in the evaluation of this last counterfactual we shall regard the last counterfactual as false. But on the Ability theory, this last counterfactual is what is expressed by the second conjunct of (3). Hence on this proposal an assertion of (3) involves the assertion of a ‘would’ counterfactual that forces one to read the ‘could’ counterfactual with negated consequent as false, and thus (given the Ability theory) to read the second conjunct of (3) as false. In other words, the assertion of the ‘would’ counterfactual that is the first conjunct of (3) forces the second conjunct to be read as the false sentence ‘ $(p \wedge q) \Box \rightarrow \blacklozenge \neg q$ ’, rather than as it appears (with only p in the antecedent), which may be true. Call this the *Forced Conflict* explanation for inescapable clashes.

The Forced Conflict explanation and the Understandable Confusion explanation are not mutually exclusive, and one suspects that they are both operative in any case of an inescapable clash on an ontic reading of the ‘might’ counterfactual. Neither is *semantically* forced, which allows (3) to be true even though it is likely to be extremely difficult to get ordinary speakers to hear the true reading of (3). This responds to DeRose’s argument against ontic theories, and serves also to respond to the supplementary argument against the ambiguity thesis, as there is now no need to posit that ‘might’ in (3) never takes the ontic reading.

These two phenomena are sufficient to explain inescapable clashes without adopting the counterfactual scepticism that an outright denial of (3) amounts to. Whatever conclusions can be drawn from the observations with which I began in §1, scepticism about almost all ordinary counterfactuals only follows if one fails to accommodate the Understandable Confusion and Forced Conflict explanations.

Earlier, when I discussed the response of the Epistemic theory to the problem of inescapable clashes I raised a further worry, which I called ‘weak counterfactual scepticism’ (§7, p. 13). No similar worry arises for the Ability theory, because knowing the truth of the ‘might’ counterfactual is not in general incompatible with knowing the conflicting ‘would’ counterfactual: there are instances of (3) that can

be both true and known true. This seems one point of advantage for the Ability theory over its Epistemic rival.

14 Ambiguity Again

In the foregoing defence of the ability theory I have mostly proceeded as if all ‘might’ counterfactuals are to be given the same semantic analysis. I now wish to question this assumption, and ask whether an ambiguity thesis is defensible for ‘might’ counterfactuals.

In fact, it seems quite clear that ‘might’, like every other modal auxiliary in English, can take on a multitude of senses. The discussion in §10 was quite explicit on this point, in its insistence that insofar as ‘might’ is intimately related to ‘possibly’, and ‘possibly’ is systematically ambiguous between epistemic and ontic senses, ‘might’ will also be infected with the same ambiguity. I think the evidence that DeRose (1991) adduces tells strongly in favour of the existence of an epistemic sense of ‘might’. I also think that the evidence I have cited above in §10, and that Stalnaker proposes (1981: 99), tells strongly in favour of a non-epistemic sense of ‘might’, one that is connected closely with ability ascriptions. The obvious profusion of senses of every other English modal auxiliary suggests strongly that ‘might’ should not be any different (Palmer, 2001). In light of the linguistic and philosophical evidence, this ambiguity between epistemic and ability readings of the ‘might’ counterfactual should be the default position.²²

Against this default position, there are a couple of responses one might make. Firstly, one could claim that ambiguity is in itself a radical response to the evidence, and that a univocal reading is preferable on methodological grounds.²³ I think one can in general agree with this line of objection, at least insofar as the methodological scruples can be made consistent with the evidence. But in this case I think the evidence, from consideration of ‘might’ along with ‘can’, ‘could’, ‘may’ and so on, tells overwhelmingly for a multitude of senses of these words. To insist on univocality in the face of this evidence on purely methodological grounds is untenable.

Secondly, one could claim that the correct reading of the epistemic ‘might’ in fact is able to account for all the apparently ontic phenomena, particularly the con-

²²This is not, therefore, the ambiguity thesis that Lewis defends, for we have seen no reason to adopt the Duality reading of the ‘might’ counterfactual as one of our potential disambiguations.

²³DeRose (1999: 402) suggests something like this objection.

nection between ‘might’ and ‘can’. As I showed earlier (§9, p. 18), a straightforward epistemic reading of ability claims is inadequate. Yet it may be that the more complicated conception of epistemic possibility developed by DeRose can explain these claims (1999: §11). In particular, the second clause of DeRose’s analysis of ‘might p ’, that ‘there is no relevant way by which members of the relevant community can come to know that $\neg p$ ’ has strongly ontic flavour to it (see §4). If there are circumstances in which the first clause of DeRose’s analysis (that ‘for all anyone knows, p ’) is presupposed or assumed, then the content of the ‘might’ claim reduces to the truth of the second clause. If we are in a circumstance in which ‘can p ’ is true, then it will often be the case that there isn’t any relevant way to find out that $\neg p$, otherwise we’d have to retract our commitment to the claim that p can be true. The main point, according to DeRose, of asserting a claim like (17), which seems to demand an ontic reading, would be to assert this second clause, and thus to assert something which is made true by the ability of the experimental apparatus to produce a certain outcome. That would then explain why, even though the ‘might’ in (17) is epistemic, it can seem to take on an ontic sense (DeRose, 1999: 402).

Though ingenious, I don’t find this argument convincing. There are a couple of problems with it. Firstly, we can utter a ‘might’ counterfactual in a clearly ontic sense even when the first clause of DeRose’s analysis is *not* presupposed:

- (42) It’s not clear whether any of our mates knows that if we had tossed the coin, it would have landed heads. Regardless, if we had tossed the coin, it might have landed tails, because it’s a fair coin.

According to DeRose, the second sentence in (42) should have an epistemic reading (it should in part communicate that neither we nor any of our mates know the ‘would’ counterfactual). But we explicitly left open the possibility that one of our mates might know that conditional; if the second sentence had the epistemic sense, (42) would be contradictory. While I agree it doesn’t sound wonderful, I doubt that (42) is bad enough to sound contradictory. On the natural ontic reading of the second clause, it is not contradictory.

Worse, there seem to be cases in which the first clause of DeRose’s analysis is known to be false, and in which a ‘might’ claim is still true, because there is a true ‘could’ claim which it expresses. Take, for example, (43):

- (43) It was fairly clear to everyone that it was actually going to turn out badly, yet because a good outcome could have come to pass everyone thought that

it still might.

In this case the explicit connection of ‘could’ and ‘might’ forces an ontic reading that can be maintained alongside the explicit disavowal of a straightforward epistemic reading of ‘might’.

Even setting aside these difficulties, I’m not at all sure that DeRose’s account isn’t just an ambiguity thesis in disguise. For his account relies centrally on the ability of the first clause of his analysis to be contextually presupposed. In these contexts, therefore, the ‘might’ claim effectively expresses the ‘could’ claim. To deny that this is an ambiguity theory requires a clear distinction between the semantics of ‘might’ and what such claims can be used pragmatically to express; I am not sure that the evidence supports such a rigid divide between the semantics and pragmatics of these expressions.

In summary, I think the evidence is overwhelming that ‘might’ counterfactuals, like ‘might’ generally, can take on at least two central senses: an epistemic sense and an ontic sense closely connected with ‘could’. The direct arguments against the Ambiguity thesis proposed by DeRose are unsuccessful.

One loose end remains; in §3, I argued against Lewis’ version of the Ambiguity thesis, and I want to reassure you that those arguments do not harm the version of the thesis defended in the present section. The first objection there was that Lewis seemed to be dragging in ambiguity simply to escape from the counterfactual scepticism that threatened his Duality account. That is not a problem for the present thesis, because as I argued in §7 and §13, both disambiguations of ‘might’ counterfactuals are sufficient to undermine counterfactual scepticism, while preserving the phenomenon of inescapable clashes. The second worry was that we had no general story to tell about when the ‘might’ counterfactual gets an ontic reading, and when it gets an epistemic reading. This worry is still somewhat pressing; yet because I’ve connected ‘might’ with ‘could’, a more general story about how and when to deploy those modal auxiliaries should shed light on our particular issue.

15 Conclusion

I’ve argued ultimately for the following claims: ‘might’ counterfactuals of the form $p \diamondrightarrow q$ are ambiguous between two different senses (§14). The first sense is given by the Epistemic theory, on which the ‘might’ counterfactual means ‘ $\diamond(p \squarerightarrow q)$ ’, where ‘ \diamond ’ is an epistemic possibility operator (§7). The second sense is given by the Ability theory, on which the ‘might’ counterfactual means ‘ $p \squarerightarrow \blacklozenge q$ ’, where

‘◆’ is the ‘can’ operator (§11). On the Ability theory, the ‘might’ counterfactual expresses the same thing as a ‘could’ counterfactual (§9). Each of these readings of the ‘might’ counterfactual is supported by considerable linguistic evidence (§10). I rejected a third approach to the ‘might’ counterfactual, Lewis’ original Duality theory, on the grounds that it did not adequately capture the semantics of ‘might’ counterfactuals as we in fact use them (§5–§6). Crucially, each of the disambiguations of ‘might’ counterfactuals allows for sentences of the form of (3) to be true while also explaining why such sentences sound invariably awful to ordinary speakers (§13). This enables us to meet the minimal conditions on the interaction of ‘might’ and ‘would’ counterfactuals without succumbing to a thoroughgoing scepticism about ordinary ‘would’ counterfactuals (§1). Of course, the phenomenon of inescapable clashes can still seem to support counterfactual scepticism. But unless we want to overturn the standard and accepted semantics for ‘could’ and ‘can’, and argue that the truth of a ‘would-be-thus-and-so’ counterfactual is in fact inconsistent with the truth of a ‘could-be-otherwise’ counterfactual, we have no direct and immediate route to such scepticism.

Faculty of Philosophy, University of Oxford
Exeter College, Oxford
antony.eagle@philosophy.ox.ac.uk

References

- Austin, J. L. (1970), “Ifs and Cans”. In *Philosophical Papers*, Oxford: Oxford University Press, pp. 205–32.
- Bennett, Jonathan (2003), *A Philosophical Guide to Conditionals*. Oxford: Oxford University Press.
- Bigelow, John and Pargetter, Robert (1990), *Science and Necessity*. Cambridge: Cambridge University Press.
- DeRose, Keith (1991), “Epistemic Possibilities”. *Philosophical Review*, vol. 100: pp. 581–605.
- (1994), “Lewis on ‘Might’ and ‘Would’ Counterfactual Conditionals”. *Canadian Journal of Philosophy*, vol. 24: pp. 413–8.
- (1999), “Can It Be That It Would Have Been Even Though It Might Not Have Been?” *Philosophical Perspectives*, vol. 13: pp. 385–413.

- Eagle, Antony (unpublished), “Probability and Ability”. In preparation.
- Edgington, Dorothy (1995), “On Conditionals”. *Mind*, vol. 104: pp. 235–329.
- Egan, Andy, Hawthorne, John and Weatherson, Brian (2005), “Epistemic Modals in Context”. In Gerhard Preyer and Georg Peter (eds.), *Contextualism in Philosophy*, Oxford: Oxford University Press, pp. 131–68.
- Fara, Michael Graff (unpublished), “The Paradox of Believability”. <http://www.princeton.edu/~7Efara/papers/believability.pdf>.
- Hawthorne, John (2005), “Chance and Counterfactuals”. *Philosophy and Phenomenological Research*, vol. LXX: pp. 396–405.
- Heller, Mark (1995), “Might-Counterfactuals and Gratuitous Differences”. *Australasian Journal of Philosophy*, vol. 73: pp. 91–101.
- Jeffrey, Richard C. (1992), “Mises Redux”. In *Probability and the Art of Judgment*, Cambridge: Cambridge University Press, pp. 192–202.
- Kratzer, Angelika (1977), “What ‘must’ and ‘can’ Must and Can Mean”. *Linguistics and Philosophy*, vol. 1: pp. 337–55.
- Lewis, David (1973), *Counterfactuals*. Oxford: Blackwell.
- (1983), “Scorekeeping in a Language Game”. In *Philosophical Papers*, vol. 1, Oxford: Oxford University Press, pp. 233–50.
- (1986a), “Counterfactual Dependence and Time’s Arrow”. In *Philosophical Papers*, vol. 2, Oxford: Oxford University Press, pp. 32–66.
- (1986b), “The Paradoxes of Time Travel”. In *Philosophical Papers*, vol. 2, Oxford: Oxford University Press, pp. 67–80.
- Mellor, D. H. (2000), “Possibility, Chance and Necessity”. *Australasian Journal of Philosophy*, vol. 78: pp. 16–27.
- Palmer, F. R. (2001), *Mood and Modality*. Cambridge: Cambridge University Press.
- Stalnaker, Robert C. (1968), “A Theory of Conditionals”. In Nicholas Rescher (ed.), *Studies in Logical Theory*, Oxford: Blackwell, pp. 98–112.

——— (1981), “A Defense of Conditional Excluded Middle”. In William L. Harper, Robert C. Stalnaker and Glenn Pearce (eds.), *Ifs*, Dordrecht: D. Reidel, pp. 87–104.

——— (1984), *Inquiry*. Cambridge, MA: MIT Press.

Williams, J. R. G. (unpublished), “Chances, Counterfactuals, and Similarity”. <http://www.personal.leeds.ac.uk/~phljrgw/wip/REVISEDcounterfactualsandchance.pdf>.