Most Counterfactuals are False

Alan Hájek

DRAFT

DEAR READER,

YOU MAY APPRECIATE SOME TIPS ON WHICH SECTIONS COULD BE SKIPPED ON A FIRST READING. IF YOU ARE LACKING IN TIME OR STAMINA, I RECOMMEND THAT YOU SKIP OR SKIM THROUGH SECTIONS 6 AND 7, SPARING YOU OVER 20 PAGES. I'VE PUT IN GREY FONT THESE, AND OTHER BITS THAT MAY BE SKIPPED IF YOU'RE IN A HURRY. FAST-FORWARDING THROUGH SECTION 5 WOULD SPARE YOU MORE PAGES, WITHOUT LOSING THE GIST OF THE PAPER.

A.H

0. Preface

Our theorizing about counterfactuals is pulled in two different directions. On the one hand, we want to make sense of the counterfactual utterances of *competent speakers*. In doing so, we strive to interpret these utterances to be true as much as we can. On the other hand, we want to make sense of the *universe*. In doing so, we strive to determine its laws, its chances, its possibilities, and other modal features of it, including the counterfactuals that it renders true—or otherwise. There is no guarantee that these two projects will harmonize. After all, competent speakers may be systematically ignorant of or mistaken about the fundamental nature of the universe, and they may insensitive to its modal features in various ways. In this book, I will argue that the universe is not as cooperative to our counterfactual discourse as we like to think: most of the counterfactuals that we utter are false. And the same goes for most of the counterfactuals that we write, believe, suppose, deliberate about, and entertain in other ways.

I have presented my arguments to many philosophical audiences over a number of years, and I tend to get rather polarized reactions. One common reaction, only slightly caricatured, is that *I have lost my philosophical marbles*—my thesis that most counterfactuals are false is so crazy that there has to be something wrong with my arguments for it. Another common reaction is that *these arguments are entirely sound and persuasive*—end of story. (Yet another common reaction is to be in an uneasy superposition of those two states: the thesis seems to be crazy, but the arguments seem to be sound and persuasive.) Along the way I have conducted an informal sociological study of how the proponents of these respective reactions break down. The first reaction comes especially from philosophers of language and mind. They tend to take as their starting point a view of semantics and of mental states according to which most of what we say or think must come out true. Here, principles of charity and other interpretivist constraints come to the fore. According to these philosophers, I must be mistaken about how our words get their meanings, or about how our mental states get their contents. The second reaction comes especially from philosophers of science and metaphysicians. They tend to agree with me that counterfactuals that involve the outcomes of chance processes are typically false, and that our best science teaches us that most of the counterfactuals in which we traffic involve just such chance processes. (Perhaps those in a superposition of states are philosophers with interests on both sides, or on neither.)

And so we have a standoff. We have the collision of two different projects: interpreting each other, and interpreting the universe. I think the first project pushes one towards the view that most of the counterfactuals that we countenance are true; the second, to the view that most of them are false. So some of my interlocutors appeal to various strategies for saving the truth of most of our counterfactuals; others agree with me that our universe is inhospitable to them, whether we like it or not.

Another possible diagnosis of the split in my audiences' reactions stems from methodological considerations about how an investigation such as this should proceed. On one view, primacy should be given to our intuitions about the *truth values* of individual sentences. They provide data points that our theories should fit as well as possible, while heeding the lesson from curve-fitting in science that we should not *overfit* the data points, since they typically suffer from some noise. This view bids us to take our prima facie judgments of which counterfactuals are true mostly at face value, and to construct our semantics for counterfactuals to respect them to a large extent. On another view, primacy should be given to our intuitions about the *validity of inferences* that we are disposed to make, or should be. This view provides no guarantee that individual counterfactuals that we utter will typically come out true—they may not if premises to which we are committed, together with inference patterns that we endorse, show otherwise.

Each of these views may need some adjusting. For our intuitions about individual sentences may not be tracking their *truth* or otherwise, but rather something else—perhaps their *acceptability*, *assentability*, *assertability*, or what have you. These '*a*'-words are meant to capture some other desirable property that these sentences might have or lack, broadly construed with another '*a*'-word of *a*pprobation: the *appropriateness* of uttering or entertaining them. And our intuitions about inferences may be not be tracking their *validity* or otherwise, but rather some other agreeable aspect of them[1]—perhaps their reasonableness as transitions in assertion or in thought, or their inductive strength.

[1] Thanks here to Daniel Nolan.

My methodology will involve the seeking of a kind of Goodmanian reflective equilibrium between our intuitions about individual sentences, and certain inference patterns in which they appear. I will begin with premises whose truth I will take to be the most secure—often because I regard them to be supported by our best science. From them I will infer the falsehood of most of the counterfactuals that we utter or entertain, by inference rules that I regard as deeply entrenched. This flies in the face of our commonsense judgment of the truth values of these counterfactuals, and thus offends the methodology that accords primacy to these judgments. However, I will explain away my insults to our intuitions: we are really intuiting another virtue of these counterfactuals, their assertability. And most of these counterfactuals are indeed assertable, their falsehood notwithstanding.

Such is my view that divides my audiences' reactions. I invite you to read on and to find out what *your* reaction is.

1. Introduction

"Even the crows on the rooftops are cawing about the question of which conditionals are true". So said Callimachus over 2000 years ago.[2] And my answer to the question is: when it comes to counterfactuals, *relatively few* of them.

The crows have much to caw about. For counterfactuals are apparently implicated in much that we hold dear. They figure in influential analyses of causation, perception, knowledge, personal identity, laws of nature, rational decision, confirmation, dispositions, free action, explanation, and so on. Science freely traffics in counterfactuals, both explicitly (in drawing consequences of its theories) and implicitly (by trafficking in concepts that are themselves tacitly counterfactual).[3] And counterfactuals are also earning their keep in the social sciences—in psychology (e.g., in understanding emotions such as regret), in history (e.g., in studying the incremental contribution by some commodity to economic growth), in the law (e.g., in apportioning responsibility), and so on.

Yet it has long been recognized that counterfactuals are strange beasts.[4] They involve a modality that makes empiricists uncomfortable; they resist truth functional analysis, yet the best-known possible worlds analyses of them make various philosophers uncomfortable; they putatively violate intuitive inference rules that material and strict conditionals obey, such as contraposition and transitivity and strengthening of the antecedent[5]; and so on. I will argue that they are even stranger than has been generally recognized: while we use them

[2] Attributed by Sextus Empiricus, Against the *Mathematicians*. Quoted in Mates (1961), 43.

[3] Consider this example from the classic physics textbook, Halliday and Resnick (19xx), p. 251: If we were to bore a hole through the earth, and drop a particle into the hole, it would move in simple harmonic motion.

[4] I have been told that in Sweden they tried to pass a law to abolish the subjunctive conditional.

[5] See Lewis 1973, §1.8. We will return to these rules in §7.

nonchalantly in daily conversation, and while they are staples of numerous philosophical analyses, most counterfactuals are false.

2. Counterfactuals under indeterminism: most ordinary counterfactuals are probably false

In what follows, by “counterfactuals” I will mean ‘would’-counterfactuals of the form ‘if X were the case, Y would be the case’, denoted ‘X □→ Y’. I will also speak of ‘might’-counterfactuals (always so-qualified) of the form ‘if X were the case, Y might be the case’, denoted ‘X ◊→ Y’. I will focus on two strategies for showing a counterfactual to be false: appealing to *indeterminism*—in particular, chanciness; and to *indeterminacy*—in particular, imprecision.

2.1 Arguments from chance

2.1.1 Chanciness Undermines Wouldiness

Stare in the face of chance. Picture as vividly as you can what a chancy process is like. For definiteness, let’s make it the toss of a fair coin. I will suppose for now that coin-tossing is a chancy business: as a coin is tossed, it is a genuinely indeterministic matter whether it lands heads or tails; and for now assume for simplicity that these are the only possible outcomes. Here is a coin that in fact will never be tossed. Consider the counterfactual:

‘If the coin were tossed, it would land heads’,

symbolically,

‘Toss □→ Heads’.

I submit that it is *false*. There is no particular way that this chancy process *would* turn out, were it to be initiated. In the words of Jeffrey (1977), “there is no telling whether the coin would have landed head up on a toss that never takes place. That’s what probability is all about”. To think that there is a fact of the matter of how the coin would land is to misunderstand chance.

Jeffrey goes on to say that asking which way the coin would land is like asking: what would be the mass of a tenth planet? There’s no fact of the matter. Indeed, one might want to go further than him on this point: it seems *worse* than asking that question. At least it’s consistent with the concept of mass that we could answer the question about the tenth planet; and perhaps cosmologists could point to facts about the distribution of nearby matter, or what have you, that would dictate a sensible answer.[6] But to say that there is a fact of the matter of how the toss would land is to deny that the coin is a chancy system, whereas the point of the example is exactly that it is.

Stalnaker (1984) taps into much the same intuition:

> Consider the following two contrasting stories:
> (1) Tweedledee and Tweedledum tossed a fair coin, but before they could see how it landed someone picked it up and ran away with it. Tweedledee is convinced that it landed heads, Tweedledum that it landed tails. Neither has any reason for his belief, but each still feels quite certain. Neither belief is justified, but one of them—we will never know which—is surely correct.
> (2) This time someone ran off with the coin before it was tossed. Having no other coin, Tweedledee and Tweedledum argue about how it would have landed if it had been flipped. Tweedledee is convinced that it would have landed heads, Tweedledum that it would have landed tails. Again, neither has a reason—they agree that the coin was a normal one and that the toss would have been fair. This time, there is

[6] Let’s ignore the inconvenient detail that astronomers *have* apparently found the tenth planet – Eris. Then by the principle of centering ((A & B) entails $A \;\Box\!\!\rightarrow B$), there *is* a fact of the matter of what the tenth planet’s mass would be: it would be the mass of Eris! But the inconvenience is minor, and it hardly scathes Jeffrey’s point. Just tweak the example: make it the *eleventh* planet. To be sure, I will later argue that counterfactuals with unspecific antecedents and comparatively specific consequents are false, so by my lights a counterfactual about the mass of the tenth/eleventh planet will presumably be false also. At this early stage, however, I think that the coin-counterfactual does seem worse.

> little inclination to say that one of them must be right. Unless there is a story to be told about a fact that renders one or the other of the counterfactuals true, we will say that neither is. (164-5)

Assuming that the toss is chancy, there is *no* story to be told about a fact that renders one or the other of the counterfactuals true. Indeed, I would go further than Stalnaker, and insist that both Tweedledee and Tweedledum are *wrong*. Imagine that they are also arguing about the colour of a bluish-green car, the former insisting that it is *blue*, the latter insisting that it is *green*, when in fact it is a borderline case, and plausibly it is indeterminate what colour it is. We may say that there is little inclination to say that one of them must be right. But somehow they are more mistaken, more deeply confused in the coin-counterfactual case than in the car-colour case. The counterfactuals that they are affirming are not merely indeterminate, the way their colour attributions are; worse than that, the counterfactuals are downright *false*.

In story (1), there is no mystery about what settles the issue of who among Tweedledee and Tweedledum is correct—there is a fact of the matter of which way the coin landed, and that's that. We may laugh at them for the firmness of their beliefs, but only because they are ignorant of this fact. Their beliefs irrationally go beyond their evidence, although we could imagine them having further evidence—e.g. the coin thief informs them which way it landed—that would *then* render one of their beliefs rational. But in story (2), their foolishness seems to be of another order. It isn't just that they are feigning knowledge of some fact of which they are ignorant; rather, there is simply no fact there to be known, no relevant evidence to be had even in principle. It is not even indeterminate whether there is such a fact, or such evidence. The hypothetical circumstances of the coin toss may counterfactually determine the *chance* of heads (the coin was specified to be fair), but *nothing* counterfactually determines how that chance process would be resolved. A further fact that would steer the process one way rather than another seems wholly mysterious—and if it existed, the process would not be chancy after all, defeating the point of the example again. So in story (2), Tweedledee and Tweedledum deserve our laughter all the more.

Our foot is in the door; now let's kick it open. While Stalnaker's stories explicitly involve a *fair* coin, in my opening coin example I deliberately did not specify the chance of heads. It is fine if you assumed that the coin was fair, but I did not need to. The argument goes through whatever the chance of tails, as long as it is a possible outcome (and our supposition of chanciness assures us that it is). For example, let the chance of tails be one in a million. Then still the counterfactual 'If the coin were tossed, it would land heads' is false. A counterfactual cannot second-guess the outcome of a chancy process, however skewed the chances are (unless the antecedent guarantees the outcome). That's what chance is all about. In a slogan: *chanciness undermines wouldiness*.

We apparently live in an indeterministic world. While that is not certain, the evidence from quantum mechanics certainly seems to point that way.[7] And it isn't just the canonical quantum mechanical examples—radioactive decay, spin measurements on a particle in a Stern-Gerlach apparatus, and so on—that are indeterministic. The indeterminism reaches medium-sized dry goods (and even oversized wet ones), just less obviously so. A cue ball colliding with an 8 ball may approximate a deterministic system, but even they are not immune from quantum mechanical indeterminism. One ball might spontaneously tunnel through the other, or to China, or to the North Star—incredibly unlikely, to be sure, but possible. Thus, I cannot truly say 'if the cue ball were to hit the 8 ball, the 8 ball would begin rolling'. Or again, whenever I jump in the air, there is a minute chance that I will not come

[7] To be sure, Bohmian mechanics is deterministic. In §4 I argue that determinism will not save most counterfactuals from being false.

down. Thus, I cannot truly say 'if I were to jump, I would come down'. With indeterminism reaching so far, a surprisingly large array of ordinary counterfactuals really have chancy consequents. Thus, they are in this sense as 'bad' as the coin tossing counterfactual with which I began. Ordinary the counterfactuals may be, but that doesn't save them from being false.

Now, there are various reasons why you may not countenance the chanciness of these processes. You may be ignorant of the relevant quantum mechanical facts. Or you may know about them, but may temporarily not be thinking about them—for example, they may simply not be relevant to your current interests, or they may not be salient in your current conversation. Or you may be thinking about them, but deliberately choosing to ignore them (the way that one can deliberately ignore an annoying TV commercial even when it is right there in front of you). For whatever reason, you may not be staring chance in the face.

The trouble is that chance is staring at *you*. Whatever you happen to know about, whatever you happen to be attending to, whatever your interests happen to be, or whatever happens to be conversationally salient, chance is constantly there in the background. It doesn't go away just because you aren't dignifying it with your thought. It is heedless of your ignorance, defiant of your ignorings.

Or you may be highly attentive to chance. For you may be a philosopher or a scientist trying to understand the universe. You may be trying to understand the nature of causation, or dispositions, or laws of nature, or other modally charged notions, and you realize that chance is intimately bound up with them. Indeed, you may be trying to understand the nature of counterfactuals. Or you may be studying what our most successful scientific theory, quantum mechanics, tells us about the universe. Once you take seriously what it says, you should see chanciness almost everywhere. But whether or not you take seriously what the theory says, that's apparently how the world *is*.

2.1.2 High chances don't save counterfactuals from falsehood

I have said that a counterfactual cannot second-guess the outcome of a chancy process, however skewed the chances are (unless the antecedent guarantees the outcome). Suppose you insist that when the chances are heavily skewed in favour of some outcome—say, 0.999999 to heads on a biased coin toss—there is really no second-*guessing* of that outcome; it *would* happen. Very well then; what do you say *would* happen on a hypothetical *second* toss of the coin? Presumably that it, too, *would* land heads. After all, the chances are heavily skewed in its favour, exactly as in the first toss. And a hypothetical *third* toss? Presumably, heads again. And so it goes. Let this sequence continue for, say, 100 million hypothetical tosses. According to you, they *would* all land heads. That's what you get when you apply even-handedly your insistence about the first toss's outcome, across the board. And you *must* apply it even-handedly, for you have absolutely no basis for singling out certain outcomes as anomalous tails outcomes—that would arguably be even more laughable than Tweedledee's and Tweedledum's unprincipled favouritism in story (2). And you have no right to claim that some tails outcomes would appear *somewhere* in the sequence, while insisting of each toss that it *would* land heads. So according to you, the coin would *never* land tails.

But that is absurd. For starters, in 100 million trials, the expected relative frequency of an event of chance one in a million is 100, not 0. Worse, *you* are now second-guessing the outcome of a chancy process *the wrong way* by your own lights—you insisted, remember, that when the chances are heavily skewed in favour of some outcome, then it *would* happen. Here are two ways the hypothetical sequence could turn out: all heads, and *not* all heads. The former has chance $(0.999999)^{100,000,000} = 3.7 \text{ x } 10^{-44}$, the latter $1 - 3.7 \text{ x } 10^{-44}$. The chances are extremely skewed in favour of the latter, and yet you are committed to the former eventuating![8]

More formally, the faulty reasoning that I have imputed to you (only as a supposition!) repeatedly appealed to the following principle:

(If High Chance Then Would)

If $p \;\Box\!\!\rightarrow (\text{chance}(q) > t)$, where t is high, then $p \;\Box\!\!\rightarrow q$.

Here, t represents a threshold above which a chance counts as heavily skewed; I assumed that 0.999999 clears this threshold. (If it doesn't, I will simply rework the example, tailoring it to the putative threshold.) On the basis of this principle, and the assumption that $t < 0.999999$, we infer from

Coin is tossed 100 million times $\Box\!\!\rightarrow$ chance(toss 1 lands heads) = 0.999999

that

Coin is tossed 100 million times $\Box\!\!\rightarrow$ toss 1 lands heads. (1)

Repeat this reasoning for the second toss, inferring from

Coin is tossed 100 million times $\Box\!\!\rightarrow$ chance(toss 2 lands heads) = 0.999999

that

Coin is tossed 100 million times $\Box\!\!\rightarrow$ toss 2 lands heads. (2)

…

And so on, through the 100 millionth toss.

The faulty reasoning also tacitly appealed repeatedly to the following principle:

[8] Thanks to Branden Fitelson for the calculation here.

(Agglomeration) If $p \,\Box\!\!\rightarrow q$ and $p \,\Box\!\!\rightarrow r$, then $p \,\Box\!\!\rightarrow (q \,\&\, r)$.

[SHOULD I STATE THIS AS AN ARGUMENT FORM RATHER THAN AS A CONDITIONAL?]

Using (Agglomeration), we can conjoin the consequents of the counterfactuals in (1), (2), and so on:

Coin is tossed 100 million times $\Box\!\!\rightarrow$

(toss 1 lands heads & toss 2 lands heads & … & toss 100 million lands heads) (*)

This is absurd on the face of it. Worse, from

Coin is tossed 100 million times $\Box\!\!\rightarrow$

chance(it does *not* land heads every time) = 1 – 3.7 x 10^{-44},

we infer by (If High Chance Then Would) again:

Coin is tossed 100 million times $\Box\!\!\rightarrow$ it does *not* land heads every toss. (**)

Applying (Agglomeration) again to (*) and (**), we infer

Coin is tossed 100 million times $\Box\!\!\rightarrow$

(toss 1 lands heads & toss 2 lands heads & … & toss 100 million lands heads

& it does *not* land heads every toss).

The antecedent is possible, while the consequent is a contradiction, so the counterfactual itself is false, indeed surely a contradiction also. Or we may appeal to Conditional Non-Contradiction, a widely agreed upon principle in counterfactual logic:

$(p \,\Box\!\!\rightarrow q) \,\&\, (p \,\Box\!\!\rightarrow \neg q)$ is a contradiction

From this we see directly that the conjunction of (*) and (**) is a contradiction.

It should be obvious that however high we set the threshold *t* in (If High Chance Then Would), I can rerun a similar argument simply by imagining a coin whose bias *p* towards heads exceeds *t*, and adjusting the number of coin tosses accordingly. I can choose the smallest *n* such that $1 - p^n > t$, and imagine the coin being tossed *n* times. On the one hand, by (If High Chance Then Would) and (Agglomeration), we have

Coin is tossed 100 million times □→ it lands heads every toss.

On the other hand, the probability that it would not land heads every toss is $1 - p^n$, which exceeds *t*. So by (If High Chance Then Would),

Coin is tossed 100 million times □→ it does *not* land heads every toss.

Something has to give: either (If High Chance Then Would), or (Agglomeration). Now, (Agglomeration) holds in Stalnaker's (1970) and Lewis's (1973) counterfactual logics, and it is highly plausible in its own right. It amounts to conjunction introduction under a counterfactual supposition. Assume that if *p* were the case, *q* would be the case, and further assume that if *p* were the case, *r* would be the case; then *q* would *still* be the case, on pain of contradicting our first assumption, so both *q* and *r* would be the case. I conclude that (If High Chance Then Would) is the culprit, however high we set the threshold *t*. *Any* chanciness in the consequent, however small, undermines the corresponding 'would' counterfactual.

As we have seen, most ordinary counterfactuals in fact have chancy consequents. It is chancy whether an 8 ball would begin rolling upon a collision from a cue ball; it is chancy whether a jumping human would come down; and so on. The fact that the relevant chances may be extremely small does not save the truth of the counterfactuals. We ought not laugh too loudly at Tweedledee and Tweedledum, for commonsense thinking about counterfactuals is guilty of a similar mistake to theirs.

2.1.3 The argument from the failure of Cournot's Principle

Plausibly, (If High Chance Then Would) is equivalent to

(If Low Chance then Would Not)

If p □→ (chance(q) < t), where t is low, then p □→ ¬q.

Here I assume that p □→ (chance(q) < t) is equivalent to p □→ (chance(¬q) > 1 – t) since their consequents are equivalent. If you think that counterfactuals with equivalent consequents need not be equivalent, then I still maintain that the two principles are equally plausible—or better, equally implausible. Note that according to these principles, no further fact is needed to drive a hypothetical chancy process one way or another when the chances are sufficiently skewed—not only are the *chances* counterfactually determined, but they in turn counterfactually determine the *outcome*. That should immediately raise alarm bells about these principles.

(If Low Chance Then Would Not) asserts that an event of low probability would not happen. It is reminiscent of the so-called *Cournot's Principle*:

> An event with low probability will not happen.[9]

While the former principle makes a set of counterfactual claims about events of low probability under a given counterfactual supposition, Cournot's Principle makes a set of outright *predictions* about the actual world. The principle has some weighty proponents, including probability theory's hall-of-famers Kolmogorov and Borel, the latter going so far as to say that it is "the only law of chance". It has currency nowadays also, Shafer being a staunch advocate.

Offhand, Cournot's Principle seems to be clearly false. It seems to be refuted every time a large lottery is played: each ticket has a low probability of winning, but *some* ticket wins.

Aristotle was on the right track when he said “It is likely that unlikely things should sometimes happen”. But I would go further: unlikely things happen *all the time*: Your reading these words, in exactly the way that you are, at exactly the time that you are, was surely antecedently improbable (in a good sense), and yet there you are doing so! Offhand it seems that every chancy event specified at a sufficient level of detail, and that every sufficiently long conjunction of chancy events, provides a counterexample to the principle.[10]

But the proponents of Cournot's Principle can fight back: they can invoke a sufficiently demanding standard for what counts as “low probability”. The million-ticket lottery that was played today does not provide a counterexample provided the bar for “low” is set below 1/1,000,000. Take the largest lottery that is ever played in the history of the universe, and whatever each ticket’s probability of winning is, make sure that the bar is set below *that*. Whatever your probability was antecedently of reading these words, in exactly the way that you are, at exactly the time that you are, make sure that the bar is set below *that*. And so on for all the putative counterexamples. Borel himself suggested that the bar could be set at 10^{-50}, but surely that is not nearly low enough. Just the conjunction of all the roulette wheel outcomes that will take place in Las Vegas tomorrow has far lower probability than that—and yet this conjunction *will* happen. Still, the thought is that there is a bar-setting that is low enough to render Cournot's Principle correct.

However, this still can’t be right, for even probability *0* events happen: indeed, they are forced upon us if we have uncountable probability spaces. Think of a radium atom decaying

[9] The principle is sometimes qualified: “An event with small probability *singled out in advance* will not happen”—see Shafer (2006). The qualification will not matter to us; just suppose that the small-probability event in question is always singled out in advance by us, or by someone who has bet on that event.

[10] To be sure, if all chances are in fact 0 or 1—perhaps because determinism is true—then Cournot's Principle is trivially true with “probability” understood as “chance”. But its proponents were surely not assuming *that*—if they had been, their statement of the principle would have been strangely coy, since

at the *exact* time that it does, where time is parameterized by a real number, and the decay law for radium is a continuous (exponential) function. Even if you are skeptical that this probability really is 0, there is still a concern. Presumably the bar for what counts as "low probability" is supposed to be set at some *positive* value. Otherwise, the statement of the principle would be quite misleading: why say that an event of *low* probability will not happen, if what was intended was that an event of *zero* probability will not happen? But if the bar is set at some value $\varepsilon > 0$, however small, then we don't need to appeal to an *exact* decay time to get our counterexample. Consider the atom decaying within a sufficiently short *interval* around its actual decay time, such that the probability of this interval is less than ε. Indeed, even making ε an infinitesimal will set the bar too high; consider an infinitesimal interval, whose probability is $\varepsilon/2$, centred on the actual decay time. Again, we have a "low probability" event that *does* happen.

So I submit that Cournot's Principle is untenable as it stands. However, something in the neighbourhood of it can be salvaged: *most* events of low probability did not happen, do not happen, and will not happen. Most lottery tickets do not win; most of the exact ways in which you might have been reading these words did not transpire; most decay times of the radium atom will not be realised. There are relatively few counterexamples to the principle; and the lower we set the bar for 'low', the fewer there will be. Every time a low probability event does not happen, the proponent of Cournot's Principle can triumphantly say: "See, I told you so!"; and he can claim *this* sort of vindication most of the time. Moreover, there is no mystery about when such vindication takes place: we just wait and see which of the principle's predictions are borne out by the actual world.

really they should have said that an event with probability *less than 1* will not happen! They surely thought

(If Low Chance Then Would Not) inherits all of the problems with Cournot's Principle and more besides, but none of its salvaging. While there is presumably the largest lottery that is ever played in the history of the universe, there is no largest *hypothetical* lottery. So we don't need to appeal to a continuous probability distribution, such as the radium decay law to create trouble for (If Low Chance Then Would Not). Wherever we set the bar (above 0), there will be a hypothetical lottery large enough to yield probabilities below that bar. It is false that most of its tickets *would not* win—more on that shortly. And the proponent of (If Low Chance Then Would Not) cannot claim the vindication of many positive instances; there is no "See, I told you so" to be had. To think that there is just is to make a version of Tweedledee and Tweedledum's mistakes.

It is tempting to say that 'if I were to jump, I would come down' is true, because the chance of my not coming down would be so low that it *would not* happen. It is tempting to say similar things in support of the truth of most ordinary counterfactuals. But the principle that events of low chance *would* not happen is even more suspect than the principle that events of low chance *will* not happen. Commonsense thinking about counterfactuals is guilty of a Cournot-like error, and worse.

2.1.4 The argument from lotteries

Let's look in more detail at counterfactuals concerning lotteries, and the further trouble they create for (If High Chance Then Would) and the Cournot-like (If Low Chance Then Would Not). They will provide us with another variant of the arguments from chanciness that most counterfactuals are false.

that probabilities could take on non-trivial values.

Set the threshold $t < 1$ for what counts as a ‘high’ probability in (If High Chance Then Would) wherever you like. (Similarly for what counts as a ‘low’ probability in (If Low Chance then Would Not). Consider a hypothetical lottery with at least $\frac{1}{1-t}$ tickets; let the number of tickets be n. We have:

Lottery is played □→ chance(ticket 1 loses) > t

and so by (If High Chance then Would),

Lottery is played □→ ticket 1 loses.

Similarly, we have

Lottery is played □→ ticket 2 loses

…

Lottery is played □→ ticket n loses.

By (Agglomeration) we have

Lottery is played □→ (ticket 1 loses & ticket 2 loses & … & ticket n loses).

But we also have

Lottery is played □→ some ticket wins.

As before, we have a contradiction. Again, (If High Chance then Would) is the culprit. (A similar argument shows (If Low Chance then Would Not) to be the culprit.)

Suppose that Tweedledee has n children, the first convinced that

Lottery is played □→ ticket 1 loses,

the second convinced that

Lottery is played □→ ticket 2 loses

and so on. They are guilty of the same irrationality as their father, just less obviously so. We know from the above reasoning, and so should they, that not all of these counterfactuals can be true; yet there is no fact that privileges any of them. And the counterfactuals are not merely indeterminate. It is not merely indeterminate whether there is such a fact; it is plainly *false* that there is such a fact.

These lessons generalize. It is false that 'If you were to play the lottery, your ticket would lose', no matter how many tickets there are in the lottery. And we need not even assume that the lottery is fair; we may heavily the rig the lottery against your ticket winning. As long as it retains any chance of winning, then we cannot truly say that it would lose if the lottery were played. After all, whatever that chance is—call it *p*—we can find another larger *fair* lottery in which *each* ticket has a smaller chance than *p* of winning. We cannot say of any ticket of the larger lottery that *it* would lose, were that lottery to be played, by our previous reasoning. So we cannot say of your ticket that *it* would lose.

In any case, fairness in the sense of equal probabilities was never the heart of the reason why lottery counterfactuals are false. Even a 'fair' lottery involves equal probabilities according to one partitioning of the outcomes, but not according to others. For example, a 'fair' million-ticket lottery can be partitioned into the outcomes 'ticket 1 wins' and 'ticket 1 does not win', with probabilities 0.000001 and 0.999999 respectively. And we have seen the folly of insisting that an event of probability 0.999999 *would* occur.

Conversely, even a heavily 'rigged' lottery can be redescribed so as to yield equal probabilities. Suppose that a million-and-one ticket lottery is highly biased, so that ticket 1 has probability 0.9 of winning, and each of the remaining million tickets have probability 0.0000001 of winning. There will be ways of subdividing 'ticket 1 wins' into 9 million cells, each of which also has probability 0.0000001. For example, place 9 million balls, numbered 1 through 9,000,000 into an urn, and choose one at random. Consider the propositions 'ticket 1 wins and ball *i* is chosen', for i = 1, 2, …, 9,000,000. The disjunction of all these propositions is equivalent to 'ticket 1 wins', but each disjunct has the same probability as each other ticket has of winning, namely 0.0000001. We dare not say that each other ticket *would not win*, in virtue of the lottery being rigged against it, for then we should say that each disjunct *would be false*—all of *these* probabilities are equal, after all. Then, by (Agglomeration), we would be forced to say that ticket 1 *would not win*, despite the heavy rigging in its favour: if each specific way of ticket 1's winning would not occur, then ticket 1's winning would not occur. A similar argument goes through however heavily the lottery is biased (falling short of certainty) towards a given ticket.

In an indeterministic world such as ours appears to be, lotteries—in a broad sense—abound. Billiard ball collisions, human jumps, and so on are lotteries in the sense that is relevant here. They are highly biased lotteries, to be sure, according to a natural subdivision of the possibilities: certain outcomes—the 8 ball rolling as expected, the jumper landing normally—are far more probable than their alternatives. But we have seen that counterfactuals concerning highly biased lotteries are false just as they are concerning fair lotteries. And there will be other legitimate subdivisions of the possibilities according to which even these highly biased lotteries are not biased. If we fine-grain sufficiently the details of a given well-behaved jump—its height falling in *this* tiny interval, its duration falling in *that* tiny interval, and so on—then we can see to it that each possibility that we countenance has equal probability, or close enough to equal probability, to that of an abnormal jump that does not end well. Indeed, if we subdivide the well-behaved outcomes finely enough, each subdivision will have *lower* probability than that of a given ill-behaved outcome. We dare not say that each subdivision of the jumper's landing *would not* occur, in virtue of its tiny probability, for by (Agglomeration) we would then be committed to saying that the jumper *would not* land. After all, the putative landing would have to occur in one or other of these finely-specified ways. But then we have no right to claim that the ill-behaved outcome *would not* occur either, for its probability of doing so is higher than each of the well-behaved outcomes.

And so it goes. Any counterfactual with a chancy consequent will go the same way, however high the chance is. But most ordinary counterfactuals have consequents that are in fact chancy. These counterfactuals are therefore false.

2.2 Arguments from 'might not' counterfactuals

2.2.1 Might-not counterfactuals undermine would-counterfactuals

Now I will change gears, turning to another route to arguing that most ordinary counterfactuals are false, this one going via corresponding 'might not' counterfactuals.

Consider the 'might' counterfactual: 'If the coin were tossed, it *might* land tails', symbolically 'Toss $\Diamond\!\!\rightarrow$ Tails'. This is true—it is implied by our supposition of chanciness of the outcome. But I claim that the original 'would' counterfactual and the 'might' counterfactual are contraries: 'Toss $\Box\!\!\rightarrow$ Heads' and 'Toss $\Diamond\!\!\rightarrow$ Tails' cannot both be true. Since the 'might'-counterfactual is true, it's the 'would'-counterfactual that must take the fall.

There are several ways to arrive at this conclusion.

The 'would/might' duality

The first way regards 'would' and 'might' counterfactuals as duals, à la Lewis (1973):

$X \Box\!\!\rightarrow Y$ is equivalent to $\neg(X \Diamond\!\!\rightarrow \neg Y)$.

Then ‘Toss □→ Heads’ and ‘Toss ◊→ ¬Heads’ are *contradictories*—they necessarily have opposite truth values. The duality of the ‘would’ and ‘might’ counterfactuals has been defended by a number of authors (e.g. Bigelow and Pargetter (1990, 103), and Bennett (2003, 192)), and assumed by others (e.g. Hawthorne 2005, and Williams 2008). It has also appealed to some authors who have written on counterfactuals in connection with the debate over whether God has ‘middle knowledge’ of the truth of counterfactuals concerning free actions—see, e.g., Adams (1977) and van Inwagen (19xx).

But my argument does not need the full strength of the ‘would/might’ duality—it suffices that ‘Toss □→ Heads’ and ‘Toss ◊→ ¬Heads’ are *contraries*.

The elimination of possibilities

Just try saying out loud:

> “If the coin were tossed, it MIGHT land tails, and it WOULD land heads if it were tossed” (*)

There is surely something defective about (*). More generally, there is serious tension between assertions of 'might’ and corresponding ‘would not’ counterfactuals, or between ‘would’ and corresponding ‘might not’ counterfactuals. DeRose (1999) calls this the phenomenon of “inescapable clashes”. Here is an argument that the clash is semantic. The ‘might’ counterfactual recognizes tails possibilities, while the ‘would’ counterfactual eliminates them. Of course, the ‘would’ counterfactual leaves open various possibilities, corresponding to various ways in which heads could be realized—the coin landing heads at noon, the coin landing heads with a soft metallic tinkle, the coin landing heads and then immediately afterwards vanishing, the coin landing heads while Dick Cheney impersonates a chicken, and so on. But they are all *heads*-possibilities; tails does not occur in any possibility left open by 'Toss □→ Heads'. On the other hand, the 'might'-counterfactual is committed to at least one tails possibility remaining live. And rightly so: clearly, 'Toss ◊→ Tails' is true. Thus, 'Toss □→ Heads' is false.

Redundancy

Suppose I say: "if I were to toss the coin, it MIGHT not land heads". I then add: "And furthermore, it's not true that it WOULD land heads if I were to toss it." You ought to be puzzled. My utterance of "furthermore" primed you to expect more information, but none was forthcoming. Instead, what came next was redundant.

Disagreement and retraction

You claim that if the coin were tossed, it would land heads. I *disagree*. One way of stating my disagreement is to remind you that if the coin were tossed, it might not land heads. We can't both be right. (In fact, how else can I *merely* deny what you said, rather than committing myself to something stronger than the denial? That is another consideration in favour of the 'would/might' duality.)

Similarly, I may assert 'If the coin were tossed, it would land heads' and later come to retract it, so that I have temporal stages of myself disagreeing with each other. One way that my later self could give grounds for his retraction would be to note that if the coin were tossed, it might land tails. (The earlier self may have had misleading evidence regarding the coin—e.g. being told by an unreliable source that it was two-headed.)

Evidence

Think of the *evidence* that I could muster in support of my claim that if I were to toss the coin, it might land tails—for example, the fact that it has landed tails in the past, or that another, similar coin that I just tossed landed tails. But arguably to the same extent, these facts counter-support the ‘would land heads’ counterfactual.

I assumed for the sake of simplicity that ‘Heads’ and ‘Tails’ were the only possible outcomes. If the assumption is false, for example, because ‘edge’ is another possible outcome, no matter. In that case, ‘Toss □→ Heads’ and ‘Toss ◊→ Tails’ are clearly not contradictories, but they are still contraries. The truth of ‘Toss ◊→ Tails’ still implies the falsehood of ‘Toss □→ Heads’. Or we may bypass considerations of ‘Tails’ altogether: with some other outcome possible, *a fortiori* ‘Toss ◊→ ¬Heads’ is true (there being a further way that heads could fail to happen), and it is contrary to ‘Toss □→ Heads’.

Our foot is in the door; now let’s kick it open. I deliberately did not specify the chance of heads. It is fine if you assumed that the coin was fair, but I did not. As before, the argument would go through equally well, whatever the chance of tails, as long as it is a possible outcome (and our supposition of chanciness assures us that it is). Then ‘Toss ◊→ Tails’ is still true, and that is all I need to establish that ‘Toss □→ Heads’ is false. For example, let the chance of tails be 0.000001. It remains true that the coin *might* land tails if it were tossed, undermining the corresponding ‘would’ counterfactual concerning heads. The point generalizes. Whenever we have a true ‘might’ counterfactual of the form *X* ◊→ *Y*, any corresponding ‘would’ counterfactual *X* □→ *Z*, where *Z* is a contrary of *Y*, is false.

The incompatibility of ‘woulds’ and corresponding ‘might nots’ underpins another argument from lotteries, this time with an extra twist at the end. Consider a hypothetical fair *n*-ticket lottery, and suppose that it's true that if the lottery were played, ticket 1 would lose. By symmetry, it must be true of each ticket *i* that if the lottery were played, ticket *i* would lose, for *i* = 1, 2, …, *n*. By (Agglomeration), we have that if the lottery were played, each ticket *would* lose. But—here is the twist—this contradicts the truth that if the lottery were played, there *might* be a winning ticket. Notice that we don't need the stronger assumption, which I made previously, that if the lottery were played, there *would* be a winning ticket. So this argument goes through even if for some reason the lottery might not yield a winning ticket.

The incompatibility of 'woulds' and 'might nots' also buttresses my claim that the relevant counterfactuals are *false*, not merely indeterminate. It is still defective, and arguably contradictory, to say:

> "If the coin were tossed, it is *true* that it MIGHT land tails, and it is *indeterminate* whether it WOULD land heads if it were tossed".

If the 'WOULD-land-heads' counterfactual is indeterminate, then surely the 'MIGHT-land-tails' counterfactual is either indeterminate or false. But the latter is not indeterminate or false; it is *true*. Determinately so: after all, it is determinately true that tails has a positive chance. Compare: "the car is blue all over" and "the car is green all over" are contraries. It is defective, and arguably contradictory, to say:

> It is *true* that the car is blue all over, and it is *indeterminate* whether the car is green all over.

Suppose it is *true* that the car is blue all over; it follows that it is *false* that it is green all over. Similarly, it is *true* that if I were to play the lottery, I might win; it follows that it is *false* that if I were to play the lottery, I would lose—no matter how many tickets there are in the lottery.

The incompatibility of 'woulds' and 'might nots' yields further results. So far I have assumed that the consequent of a given *would*-counterfactual has chance less than 1 in order to guarantee that that the corresponding *might not*-counterfactual is true, and thus to show that the *would*-counterfactual is false. But we can even drive the chance of the consequent all the way to 1 and *still* have a false *would*-counterfactual. Now suppose that the coin is to be tossed repeatedly infinitely many times. It *might* land tails on every toss, even though the chance of this is 0.[11] (Again, I deliberately did not specify the chance of heads: as long as tails is a possible outcome of any toss, it is a possible outcome of *every* toss, assuming that the trials are independent.) Thus, I cannot truly say 'if I were to toss this coin forever, it would land heads eventually'. It might not.

This infinite coin toss experiment corresponds to an infinite, highly biased lottery: ticket #*i* wins iff the coin lands heads for the first time on the i^{th} toss. In this lottery, it is possible that no ticket will win (even though this can only happen in one way, and it can fail to happen in infinitely many ways—to which, moreover, the probability distribution is extremely heavily biased if the coin is fair). So I cannot even truly say 'if I were to hold all the tickets in the lottery, I *would* win'. I *might* not.

Returning to more ordinary counterfactuals: I cannot truly say 'if the cue ball were to collide with the 8 ball, the 8 ball *would* begin rolling'. It *might* not: the cue ball might quantum tunnel to China. I cannot truly say 'if I were to jump, I *would* come down'. I *might* not: I too might quantum tunnel to China! And so it goes.

At this point you may want to protest.

> PROTEST: Granted, anomalous results such as a billiard ball quantum tunneling to China or a jumping person failing to land have positive chance of occurring. But in the *nearest*

[11] See Williamson (2007) for an argument that this probability is not infinitesimal—it really is 0.

> possible worlds in which the relevant antecedents are true, these outcomes do not occur. A person's quantum tunneling mid-jump, for example, is such a *bizarre* event that a world in which it happens is rather remote from ours, and in particular more remote than worlds in which he falls normally. And we should adopt a Stalnaker/Lewis-style semantics for counterfactuals: roughly, $p \Box\!\!\rightarrow q$ is true just in case at the nearest worlds where p is true, q is true. Thus, the counterfactuals come out true after all: 'if the cue ball were to collide with the 8 ball, the 8 ball would start rolling', 'if I were to jump, I would come down', and so on.

I believe that you have effectively denied that the 'might'-counterfactuals are true. For example, in claiming that all the nearest worlds in which I jump are worlds in which I fall normally, you have ruled out my quantum tunneling from being among them, thus apparently ruling out that I might quantum tunnel were I to jump. But what entitles you to do that? Not its small chance—recall my discussion of Cournot's Principle.

I put the word "*bizarre*" in your mouth during your imagined Protest, and presumably that goes beyond merely "having small probability". After all, your reading these words, exactly as you are, was antecedently improbable, but it is not (I hope!) a bizarre event. Fleshing out what "bizarre" really means will be no mean feat. Still, I agree that in some good, intuitive sense, there is something strange about the events I am imagining. My ill-fated jump's positive probability does not save it from that epithet. But I am asking very little of that ill fate: all I need to be true is the modal claim that *I might* quantum tunnel, were I to jump. If you want to take issue with that, you have not only me but also quantum mechanics to contend with.

So if nothing else, quantum mechanics is there to guarantee the truth of the 'might' counterfactuals that undermine the corresponding 'would' counterfactuals. But often less esoteric facts will equally do the job. Holding my cup tantalizingly over a hard floor, you say: "If I were to release the cup, it would fall. And if it were to fall and hit the floor, it would break." Well, no, and no—it might not, and it might not. If I were to release the cup, a sudden gust of wind might lift it higher; and if it were to fall and hit the floor, another gust of wind might slow down its fall sufficiently to spare it a damaging impact. Or even less esoterically, I might catch the cup, sparing it an impact altogether. Quantum mechanics is just a handy, cover-all way for me to secure the truth of a huge raft of undermining 'might not' counterfactuals in one fell swoop. But other anomalous happenings could do the job just as well on a case-by-case basis.

2.2.2 Quasi-miracles

You may grant me that that there would be a *chance* of these anomalous happenings, but still deny that they *might* happen if the relevant antecedents were realized. You remind me that your protest did not question the chanciness of such bizarre outcomes, but rather their occurrence in the *nearest* worlds where the antecedents are true—their bizarreness relegates them to more distant worlds. Lewis (1986) would call them *quasi-miracles*. He characterizes the notion of a quasi-miracle with these phrases: "a remarkable coincidence … the remarkable way in which the chance outcomes seem to conspire to produce a pattern" (60). It is an event that, while compatible with the laws in virtue of its chanciness, detracts from overall similarity to the actual world in virtue of its remarkableness.

I find the notion of a quasi-miracle problematically unclear, given how much theoretical weight Lewis wants to place on it. Moreover, it *is* clear that it crucially involves the idea of being "remarkable". But this sounds suspiciously anthropocentric, a matter of what *we* find especially surprising. Whether or not something "seems to conspire" to produce pattern sounds similarly anthropocentric. One wonders how there could be a place for such notions in an objective scientific theory of the world, the way that there apparently *is* a place for the notion of chance. Ironically, just a page earlier Lewis chides (a standard version of) quantum mechanics for its "anthropocentric foundation" in the projection postulate, which says that measurement reduces the wave function. My concern is that he is giving his theory of counterfactuals such a foundation. And since he believes that causation, dispositions, and persistence through time are all ultimately to be analyzed in terms of counterfactuals, these too seem to have such a foundation, by his lights.

Be that as it may, seen from the perspective of chancy laws, there should be no prejudice against the anomalous events that I am countenancing. They should not, as it were, be penalized twice—once for their low chance, and a second time for their quasi-miraculousness. Rather, the laws treat them even-handedly, much as we treat lottery outcomes—sensitive to their low chance, but no more. After all, 'chances' are written into the laws themselves, but 'degrees of quasi-miraculousness' are not.

But if we insist that quasi-miracles send us to more distant worlds, then note that we seem to get the result that *nothing* remarkable happens in our neighbourhood of possible world space. For example, let's agree that it is remarkable if a fair coin that is tossed a 100 times in a row lands all heads. Then we seem to have the result that even if the 100-toss experiment were run a googolplex times, in *none* of the experiments would we get all heads; after all, for each experiment, *its* yielding the remarkable all-heads outcome detracts from overall similarity to the actual world. ((Agglomeration) formalizes this reasoning.) More generally, it would seem that our neighbourhood of worlds, despite its populousness, consists solely of *dull* worlds, ones free of any remarkable coincidences that are absent in our world. (I am indebted here to Hawthorne 2005.) Indeed, a sufficiently rich world in which *nothing* remarkable occurs seems rather remarkable for that very reason! Imagine a world with billions of long runs of coin tosses, trillions of people repeatedly playing lotteries, zillions of monkeys hitting typewriters at random … and yet none of them produces a remarkable run of outcomes. How remarkable! The absence of a quasi-miracle at one level itself gives rise to a quasi-miracle at the meta-level. One wonders whether the idea of a sufficiently rich world entirely free of quasi-miracles is even coherent.

Williams (2007) offers an account of quasi-miracles in terms of a notion of 'typicality'. While each possible sequence of coin tosses is equally probable, well-mixed sequences are *typical*, while a sequence of all heads is *atypical*. (He adopts an analysis of 'typicality' given by Elga 19xx.) Quasi-miracles, then, can be regarded as involving *atypical* patterns, and they detract from similarity: worlds in which they occur are *ipso facto* rendered less similar than worlds in which they do not. It is an attractive idea, but I believe it yields unattractive results. According to it, it is an analytic truth that if a coin were to be tossed many times, it would yield a *typical* sequence of outcomes. After all, according to the view, it follows from the semantics for counterfactuals that all the closest possible worlds in which the coin is tossed many times are ones in which it behaves typically. But I maintain that if a coin were to be tossed many times, it *might not* yield a typical sequence. Far from the 'would' counterfactual being an analytic truth, I deny that it is true at all. But all I need for my argument against Williams' proposal is that it is not an analytic truth.

2.2.3 Problems with the similarity semantics for counterfactuals

Yet I am still being too concessive to your Protest. You have insisted that my mid-jump tunneling takes us to *less similar* worlds than worlds where things go as expected. Let me grant that for the sake of the argument. It is still a further step for you to reach the conclusion that the counterfactual ‘if I were to jump, I would come down’ is *true*. “Ah, but that follows from the usual Stalnaker-Lewis style semantics for counterfactuals”, you say. Agreed—but perhaps we should question such similarity-based semantics. In the thick of a disagreement about the orthodoxy concerning the truth-values of counterfactuals, to presuppose the orthodoxy of some similarity-based semantics for counterfactuals is to put the cart before the horse. So let us revisit that orthodoxy. I will now argue that the connection between similarity of worlds and the truth-conditions for counterfactuals is far less straightforward than has been widely assumed.

A chancy coin is hooked up to a Doomsday machine. If the coin is tossed and it lands heads, nothing interesting happens: it’s business as usual, status quo. If the coin is tossed and it lands tails, something very interesting happens: the Doomsday machine obliterates the world and surrounding districts, resulting in vast, widespread changes to the status quo. In fact the coin is never tossed. But what would happen if it *were* tossed? Bad answer: it *would* land heads. Bad answer, because the chanciness of the coin should preclude us from giving this verdict. The chanciness of the coin to be incompatible with the truth of ‘Toss □→ Heads’. Nevertheless, various similarity accounts appear to be committed to the truth of this counterfactual. After all, intuitively the nearest ‘Toss’ worlds are ‘Heads’ worlds—business as usual is more similar to the actual world than Doomsday.

This example evokes Fine’s (1977) famous argument against Lewis: it appeared that Lewis was committed to denying that ‘if Nixon had pressed the button, there would have been a nuclear holocaust’. Lewis, equally famously, replied by laying down an ordering of what matters in judgments of similarity of worlds: “(1) It is of first importance to avoid big, widespread, diverse violations of law. … (4) It is of little or no importance to secure approximate similarity of particular fact" (1986, 47-48). I elide over the details of Lewis's ordering. The upshot was that he argued that this ordering vindicated the intuitively correct verdict on the Nixon counterfactual.

My example differs from Fine's in being *indeterministic*. The chanciness of my coin should bar us from judging 'Toss □→ Heads' to be true—that would be tantamount to second-guessing an avowedly indeterministic process. Fine's example also turns partly on intuitions about closeness of miraculous reconvergence worlds, which Lewis's ordering was meant to undermine. My example requires no miracles; if it matters, add chanciness at every stage of the relevant causal chains. I merely require the seemingly unassailable judgment that a business-as-usual world is more similar to actuality than a Doomsday-world.

Now, Lewis offers his initial ordering for *deterministic* worlds. He goes on to consider an indeterministic version of Fine's case, but the upshot is the same: "still we can say that [approximate convergence of particular fact after Nixon's button-pressing] counts for little or nothing, so it is not so that if Nixon had pressed the button there would have been approximate convergence to our world, and no holocaust" (60-61). He would clearly say the same about my example. After all, he can argue that the Heads world is not *really* business-as-usual. There will be various traces of the coin landing Heads that are absent in the actual world, so at best the Heads world is only approximately similar to the actual world regarding particular fact; and approximate similarity of particular fact "counts for little or nothing". But then he had better harden this line: approximate similarity of particular fact had better count for *nothing*. Otherwise, I will insist that the approximately-matching Heads world is at least *a little* more similar to actuality than the Doomsday world—offhand it surely seems that way—

and since comparative similarity is all that matters to the truth-conditions of counterfactuals, that's good enough to secure the truth of 'Toss □→ Heads'. In other words, to respect my insistence that we must not judge 'Toss □→ Heads' to be true, we must judge the approximately-matching Heads world to be no more similar *at all* to actuality than the radically-non-matching Doomsday world. But surely that does violence to our intuitions about similarity. And once we do that, there is a danger of losing the intuitively correct verdicts about ordinary counterfactuals in any case. Recall that your imagined protest took it for granted that the nearest worlds to ours are ones in which billiard ball collisions or human jumps transpire much as we expect them to. If approximate similarity of particular fact counts for *nothing,* then all bets are off.

In the indeterministic case, we may be able to achieve *perfect* match of particular fact after the time of the antecedent. Lewis goes on to consider the case in which thanks to a pattern of outcomes of chance processes, perfect convergence is achieved—e.g. all traces of Nixon's button-pushing are entirely obliterated. But he insists that such convergence is a *quasi-miracle*, and therefore will not occur in the nearest worlds where Nixon pushes the button. However, I have voiced my skepticism about appealing to the notion of a quasi-miracle in support of claims about the truth of counterfactuals.

In any case, we can cut through much of this intuition-mongering about similarity of worlds by using the following recipe for generating counterexamples to at least some similarity-based accounts of counterfactuals. Let the proponent of such an account go first: tell us your ordering for similarity (much as Lewis did). I will then attempt to fashion a coin case accordingly: the scenario resulting from Heads is concocted to be something judged more similar *by that ordering* than the scenario resulting from Tails. The business-as-usual/Doomsday scenarios merely made vivid the recipe for a plausible-looking similarity judgment.

Now, you may be able to thwart me by building criteria into your similarity ordering that prevent me from hooking up Heads to a more similar scenario, Tails to a less similar scenario. You may no longer be so concerned to give a *similarity* account. After all, the intuition that 'business *almost* as usual' is more similar than 'Doomsday' takes a lot of explaining away.

For example, you might say that all that matters to similarity is the *past* relative to the antecedent, and the coin's landing one way or another cannot affect that. (See Jackson 19xx.) It would be less misleading to call this a similarity of *pasts* account, rather than of total worlds. Then future similarity—indeed, even perfect match—counts for *nothing* for you. Whether business is almost as usual, or totally unusual, after the coin toss makes no difference on this view, because that's all *future* business. But then you would seem to have abandoned all resources for adjudicating counterfactuals with consequents that lie in the future relative to their antecedents. There's no saying what the billiard balls, or the jumping human, would do in even the near future.[12]

So perhaps you claim that 'similarity' is a purely technical notion, a relation that simply induces an ordering on worlds suitable for the Stalnaker-Lewis style evaluation of counterfactuals. That flies in the face of the vast bulk of literature regarding 'similarity' accounts, which has been driven by intuitions about *similarity* and not something else—to be sure, a somewhat refined notion of 'similarity'—and which has apparently found them to be

[12] Or perhaps your similarity account is antecedent-relative, building in a dependence of the similarity relation on the antecedent itself (cf. Kment 200x, Schaffer 200x). Then it would be a similarity of worlds-*relative-to-antecedents* account, rather than a similarity of worlds account per se. Let's not quibble about names, and let's avoid a lengthy digression into the pros and cons of this alternative to the usual Stalnaker-Lewis style semantics for counterfactuals. I merely want to stress that a quick invocation of that semantics in support of the imagined protest is surely *too* quick.

of heuristic value. If you are breaking free of that literature, again you would do better to give your relation a less misleading name, one with no prior associations—say, 'the R-relation'—in order to forestall misunderstandings. Then presumably we are not supposed to have intuitions one way or another about it, except by reverse engineering, working backwards from the counterfactuals that we think are true to what the R-relation would have to be to deliver those verdicts. Clearly it would be question-begging to appeal to this R-relation to support your protest, when the truth-values of such counterfactuals are the nub of our disagreement. And your protest was plausible only to the extent that an *intuitive* notion of similarity was assumed—you insisted, remember, that the *most similar worlds* were ones in which bizarre things did not happen, and you *inferred* from this that the ordinary counterfactuals about billiard balls and jumps must be *true*. But if 'similarity' is some purely technical notion, again all bets are off.

In any case, whether the relation is recognizably a 'similarity' relation or not, if it permits me to use my recipe, all is lost for an account of counterfactuals based on it. As long as I can fashion a case in which heads yields a closer world, tails a less close world *according to the ordering induced by R,* the account will predict that 'Toss □→ Heads' is true—an unacceptable result.

Our foot is in the door; now let's kick it open. I deliberately did not specify the chance of Heads. It is fine if you assumed that the coin was fair, but I did not. The argument would go through equally well, whatever the chance of Heads, as long as it is a possible outcome (and our supposition of chanciness assures us that it is). Now make the coin highly biased to tails. Still, I will tailor my example to your similarity ordering, so that heads results in a more similar scenario by your lights. Then despite the bias to tails, you will be committed to affirming that if we were to toss the coin, it would land heads. Indeed, we can drive the chance of one counterfactual's consequent all the way to 1 and *still* have you affirm the *other* counterfactual. We set up an infinite sequence of tosses, such that if there is *ever* a tail, the resulting scenario is more dissimilar to actuality than if there is *never* a tail. Then you must affirm that if we were to toss the coin infinitely many times, it would land heads every time. Similarity theorists of counterfactuals: beware![13] This completes my reply to your Protest.

And so I reach the interim conclusion that *most ordinary counterfactuals are probably false*. There are three weasel words here: '*most*', '*ordinary*', and '*probably*', and soon I will strip away two of them—the title of this manuscript, after all, is "Most Counterfactuals are False" without further qualification. "*Most*" remains, because I will eventually concede that *some* counterfactuals are true. But it will turn out that they are typically *extraordinary*—rarified, recondite, recherché counterfactuals that philosophers may occasionally traffic in, but not normal people. In the meantime, I can do away with the qualifier 'ordinary'. Time for some stripping.

3. Most counterfactuals are probably false

I have spoken of "*ordinary*" counterfactuals, but I don't want to fuss much about *defining* when a counterfactual is ordinary. (Just try, if you think it's easy!) Let me generously count as ordinary pretty much any counterfactual that you hear uttered on the street, or indeed outside a philosophical discussion. I'm being generous, because I'm prepared to count as ordinary many an arcane counterfactual from science—some biochemist's counterfactual about a reaction rate, some astrophysicist's counterfactual about a galaxy red shifting, and so on. If

[13] My homage here to Lewis's dust-jacket tribute to David Stove's *The Plato Cult* is intentional.

you like, understand the claim that ‘most counterfactuals are ordinary’ so that it is an analytic truth (replace ‘ordinary’ by ‘typical’ if that helps).

Indeed, we will see in §6.3 that even large classes of *extraordinary* counterfactuals that appear to be true may not be so, thus swelling the ranks of the false counterfactuals still further. I will not attempt a census of just what proportion of all counterfactuals the true ones constitute—an impossible task—but I think that it will be *clear* that they are in the minority, and a small minority at that.

You may be tempted to say that there are uncountably many counterfactuals, of which uncountably many are true, in which case it makes no sense to speak of ‘proportion’, ‘minority’, and so on without some measure defined over sets of counterfactuals. When I quantify over counterfactuals, I don’t mean all possible counterfactuals, the overwhelming majority of which could not even be asserted in a human lifetime because their antecedents and consequents are so complex. I mean instead the counterfactuals that one hears and reads in daily life. Imagine, if you like, a transcript of all counterfactuals ever uttered or written in the whole of human history, past, present and future. Needless to say, the set of all such counterfactuals is finite. And the vast majority of *them,* I am arguing, are probably false.

I say *“probably”* false, because so far I have assumed that the world that we live in is indeterministic. While I am not certain that it is, I think it is reasonable to be fairly confident that it is. My argument exploited possible ways in which things might turn out differently from what various ‘would’ counterfactuals claim, and so far I have appealed to indeterminism to guarantee that there are such possibilities. In a deterministic world, I risk losing that argument. But determinism is probably false.

No matter—determinism will restore truth to few of our counterfactuals in any case. Time for more stripping.

4. Counterfactuals under determinism: most counterfactuals are false

4.1 Arguments from indeterminacy

4.1.1 Indeterminacy undermines wouldiness

Even determinism will not save most counterfactuals from falsehood. A number of authors argue that determinism will not save the world from chanciness—"compatibilists" about chance such as Arntzenius (200x), Eagle (200x), Hoefer (forthcoming), Levi (19xx), and Loewer (2001). If they are right, it would seem that I do not lose my argument from chanciness to the falsehood of counterfactuals after all. But let's assume they are wrong, if only to make my job harder. I don't need the chanciness strategy for arguing for the falsehood of counterfactuals, because I have another one—the one that goes via considerations of *indeterminacy*.

For even if our world is deterministic, in the neighborhood of any trajectory of an object (of a billiard ball, or of a jumping human, or what have you) there will typically be some extraordinary trajectory in which things go awry. Here I appeal to statistical mechanics, whose underpinnings are deterministic. The point is familiar from the diffusion of gases, made vivid by Maxwell's demon. (Much as it is fair game for the epistemologist to remind us of the evil demon, it is fair game for me to remind us of Maxwell's demon.) For every set of initial conditions in which the air molecules in my office remain nicely and life-sustainingly spread throughout the room, there is a nearby initial condition in which they deterministically move to a tiny region in one corner—'nearby' as determined by a natural metric on the relevant phase space. So it is false to say that if I were in my office, I would be breathing normally; the initial conditions *might* be unfortunate ones for me, leading to a phase-space trajectory of the molecules that suffocates me. The point generalizes to other deterministic systems. For every set of initial conditions in which the cue ball hits the 8 ball and each follows an expected trajectory, there is a nearby initial condition in which the balls behave anomalously. For every set of initial conditions in which I jump and land normally, there is a nearby initial condition in which I vaporize instead.

Now I exploit not indeterminism, as I did previously, but rather *indeterminacy*—in particular, the sort of indeterminacy that is due to *imprecision*, or *underspecification*. The antecedent of "if I were to jump, I would come down" is imprecise: I have not told you anything about the manner in which my hypothetical jump takes place, let alone given you a molecule-by-molecule specification of the jump. The antecedent, then, covers a huge range of initial conditions, each of which results in my jumping. Among them will be initial conditions that give rise to anomalous trajectories in which I vaporize, for the antecedent is too imprecise to rule them out. To be sure, the anomalous trajectories are sparse among all possible trajectories. But they exist all the same, compatible with the imprecisely specified conditions given in the antecedent.

As in the argument from indeterminism, I think this argument is best deployed directly. A counterfactual cannot second guess the resolution of an indeterminacy. In a slogan: *indeterminacy undermines wouldiness*.

But let's also cast the argument in terms of 'might not' counterfactuals. If I were to jump, I might wind up on one of those anomalous trajectories. Thus, it is false to say that if I were to jump, I would come down. I might not.

Now you may want to lodge your second protest.

> SECOND PROTEST: Granted, these anomalous trajectories are compatible with the antecedent, but the *nearest* worlds in which the hypothetical jump takes place are ones in which you come down. So it is true after all that you *would* come down were you to jump.

This is similar to your previous protest, and my reply is similar; it's just that now it is statistical mechanics rather than quantum mechanics that you are taking on. For you are denying that things *might* go anomalously in the ways I have imagined, while one of the most interesting features of statistical mechanics is to assert just that.

Or perhaps you are merely driving my argument in reverse, *tollensing* where I *ponensed*. You take as your starting point the apparent platitude that if I were to jump, I *would* come down. I am assuming you to agree with me that 'would' and 'might not' counterfactuals are contraries (we will drop this assumption in §§5.3 and 5.4). So you deny that if I were to jump, I *might not* come down. To be sure, they say that "one person's modus ponens is another person's modus tollens". But I claim to have physics on my side.

As before, often less esoteric facts will do the job of securing the truth of the 'might not' counterfactuals. If were to jump, a huge gust of wind might lift me higher. Statistical mechanics is just a handy, cover-all way for me to secure the truth of a huge raft of undermining 'might' counterfactuals in one fell swoop. But other anomalous happenings could do the job just as well on a case-by-case basis.

4.1.2 More problems with the similarity semantics

Yet I am still being too concessive to your Second Protest. As in your First Protest, you have insisted that my mid-jump vaporization takes us to *less similar* worlds than worlds where things go as expected. Again, let me grant that for the sake of the argument. Again, it is still a further step for you to reach the conclusion that the counterfactual 'if I were to jump, I would come down' is *true*.

For now consider the problems that imprecision or underspecification create for at least some similarity-based semantics for counterfactuals. Consider the counterfactual 'if I were at least 7 feet tall, I would be *precisely* 7 feet tall (precise to infinitely many decimal places)'. I hope you agree with me that this is *false:* it seems absurd to affirm a counterfactual with such an imprecisely specified antecedent, and yet such a precisely specified consequent. Or perhaps you think that it is *indeterminate*, but you still agree with me that it is *not true*. Yet Lewis for one is apparently committed to it being *true*.

This example evokes Lewis's (1973) famous argument against Stalnaker's *limit* assumption[14]: that for any possible antecedent *X*, there is at least one nearest *X*-world. Lewis challenged this by considering counterfactuals of the form 'if I were over 7 feet tall, then …' What are the nearest worlds where the antecedent is realized? Try to pick such a world—say, one in which I am 7 feet 1 inch tall. Surely a world in which I am 7 feet ½ inch tall is closer to actuality—after all, in that world I am closer to my actual height. But a world in which I am 7 feet ¼ inch tall is closer still to actuality; and so on, ad infinitum.

Fair enough; but with the tiny tweak that I have given it, the example backfires on Lewis. Changing the antecedent from 'I am over 7 feet tall' to 'I am *at least* 7 feet tall' gives Lewis no wiggle-room—he is apparently committed to the most similar such worlds being those in which I am *exactly* 7 feet tall. After all, those are the 'I am at least 7 feet tall'-worlds in which I am closest to my actual height.[15]

[14] And also an argument by Pollock (1976) against Lewis's denial of that assumption.

[15] You may wonder whether atomism or genetics may cast some doubt on this. Maybe my exceeding 7 ft tall by sub-atomic distances does not detract from similarity; moreover, maybe facts about genetics could make my nomically possible heights 'granular', with some slight overshooting of 7 ft possible, but any

Lewis is *apparently* so committed—perhaps his example is only supposed to illustrate the possibility of a kind of structure that is problematic for Stalnaker's theory, and we shouldn't take this particular example too seriously. In that case it would be nice to see an example that we *should* take seriously, so that we are convinced that Lewis's concern is not merely a theoretical possibility, that it actually arises for counterfactuals that we might assert or believe. I wager that I could rewrite my objection, mutatis mutandis, using any such example. In the meantime, I believe that appealing to Lewis's own example is a fair *ad hominem*.

This suggests the following recipe for generating counterexamples to at least some similarity-based accounts of counterfactuals. Let the proponent of such an account go first: tell us your ordering for similarity. I will then attempt to fashion a counterexample accordingly: a counterfactual with a highly imprecise antecedent, and a comparatively precise consequent, such that all the closest worlds *by that ordering* that realize the antecedent realize the consequent. The danger is that you will then be committed to affirming the counterfactual, whereas we should baulk at its highly unspecific antecedent coupled with a comparatively specific consequent. Once again, similarity theorists of counterfactuals: beware!

This completes my reply to your Second Protest.

4.2 The argument from disjunctive antecedents

I turn to a closely related argument, which again I take to show the falsehood of counterfactuals whether or not determinism is true. I will call it *the argument from disjunctive antecedents*.

smaller overshooting not. Of course, these would also be objections to Lewis's original argument against Stalnaker. In any case, we may circumvent them with a minor revision to the example. Consider a line in an atomless world that is 6 ft long, and now entertain counterfactuals about what would be the case if the line were at least 7 ft long.

What do you think of this counterfactual: ‘If we were in New York or Baghdad, then we would be in in the U.S.’? I submit that it is *false*, and I hope you agree. There is a way of realizing the antecedent—the Baghdad way—according to which we would not be in the U.S..

Now add some more American disjuncts, but keep the Baghdad disjunct:

‘If we were in New York, or Los Angeles, or San Francisco, or Chicago, or Baghdad, then we would be in the U.S..’

I say that this is still false. You can’t save the counterfactual by adding more ‘good’ disjuncts. Just one bad one poisons the whole counterfactual. Let’s call this *the poisoning principle*:

The disjunctive-antecedent counterfactual

$(D_1$ or D_2 or … or $D_n) \Box\!\!\rightarrow C$ is false

if any of the individual-disjunct counterfactuals

$D_i \Box\!\!\rightarrow C$ is false.[16]

Now consider your favorite counterfactual that you regard as true—let’s suppose that it is ‘if I were to jump, I would come down’. I can replace the antecedent with a long disjunction of all the specific ways of my realizing a jump:

‘if I were to jump according to initial conditions x,

or

according to initial conditions y,

or

…

or according to initial conditions z,

[16] This is closely related to the so-called “Nute’s Principle” (offered by Nute 1975).

I would come down'.

I will be able to come up with a poisoning disjunct—say, a Maxwell's-demon-style initial condition according to which I would *not* come down. It is like a 'Baghdad' disjunct. It will poison the whole counterfactual.

In fact, I don't even need the poisoning disjunct to be one according to which I would *not* come down. It suffices for it to be one for which it is false that I *would* come down. For example, suppose that according to disjunct D_i it is *chancy* whether or not I come down. Then, I say,

$D_i \Box\!\!\rightarrow$ I come down

is false, although so is

$D_i \Box\!\!\rightarrow$ I do not come down.

If D_i *leaves it open* whether or not C would be the case, then

$D_i \Box\!\!\rightarrow C$ is false.

This is plausible in its own right, but we may argue for it via a 'might-not' counterfactual. If D_i *leaves it open* whether or not C would be the case, then if D_i were the case, C *might not* be the case:

$D_i \Diamond\!\!\rightarrow \neg C$ is true.

From this it follows that

$D_i \Box\!\!\rightarrow C$ is false,

and by the Poisoning Principle that is sufficient for the falsity of

$(D_1$ or D_2or … or D_n) $\Box\!\!\rightarrow C$.

Conditional non-contradiction guarantees that

$D_i \Box\!\!\rightarrow \neg C$

is sufficient for

$$\neg(D_i \Box\!\!\rightarrow \neg C).$$

But the former is not *necessary* for the latter unless conditional excluded middle holds:

$$(X \Box\!\!\rightarrow Y) \vee (X \Box\!\!\rightarrow \neg Y) \text{ is a tautology.}$$

Indeed, I deny that it holds. Cases of chanciness provide counterexamples, in my opinion: both

$$\text{Toss} \Box\!\!\rightarrow \text{Heads}$$

and

$$\text{Toss} \Box\!\!\rightarrow \neg\text{Heads}$$

are false, and thus so is their disjunction. My opposition to conditional excluded middle should come as no surprise in any case. According to it, the negation of a counterfactual is equivalent to another counterfactual: $\neg(X \Box\!\!\rightarrow Y)$ is equivalent to $(X \Box\!\!\rightarrow \neg Y)$. But according to me, most counterfactuals are false, and so the negations of most counterfactuals are true. These negations had better not be equivalent to counterfactuals, for most of *them* I will regard as false also!

In sum, 'If we were in New York or Baghdad, then we would be in in the U.S.' is false, because according to the 'Baghdad' disjunct, we *would not* be in the U.S.: the disjunct *guarantees* that the consequent is false. But a disjunct doesn't have to be as poisonous as that to poison a disjunctive-antecedent counterfactual in which it appears. It suffices for the disjunct to *leave it open* whether or not the consequent would be realized; it suffices that the consequent *might not* be realized, were the disjunct to be realized. This makes poisoning a counterfactual surprisingly easy.

The poisoning principle spells still more trouble for similarity accounts of counterfactuals. (Again, Fine has given a similar argument.) Let us agree that a world in which we are in New York is more similar than one in which we are in Baghdad—it would require more of a change in our situation (e.g. our beliefs and desires) to get us to go to Baghdad than it would to get us to go to New York. (Change 'Baghdad' to 'Alpha Centauri' if you're not convinced.) Then the most similar worlds in which the antecedent 'we are in New York or Baghdad' is realized are ones in which we are in New York; and in those worlds, we *are* in the U.S.. So similarity accounts should judge the counterfactual to be true. I submit that this judgment is wrong.

* * * * * *

So not even determinism suffices to save counterfactuals from falsehood. I said that determinism is probably false, but it turns out it doesn't matter either way. Whichever way our world goes, *most counterfactuals are false*.

5. Fall-back/rival positions

[NOTE TO READER: THIS SECTION MOSTLY CONSIDERS RESPONSES TO THE 'MIGHT NOT-COUNTERFACTUALS UNDERMINE WOULD-COUNTERFACTUALS' ARGUMENTS. IN A FUTURE DRAFT I WILL DISCUSS IN MORE DETAIL RESPONSES TO THE ARGUMENTS FROM CHANCE AND FROM DISJUNCTIVE ANTECEDENTS.]

Philosophy, like tight-rope walking, is a risky business.[17] And a philosopher, like a tight-rope walker, is well advised to have safety nets in place—especially when arguing for a

[17] Compare the opening words of Lewis (1969).

position that is likely to be controversial, as I am. So in the next few sections I want to offer some fall-back positions—although they are rivals to my view, and I will not retreat them without a fight. I have given you arguments for a radical position, but if you are unconvinced by them, you may still be convinced by these relatives of them. These relatives have weaker conclusions, but I think they are still radical enough to be interesting.

5.1 The counterfactuals' truth values are context-dependent rather than false

You may say that the counterfactuals that I have countenanced are not uniformly false, but rather have *context-dependent* truth-value. For example, in normal conversational contexts, it is true to say 'if I were to jump, I would come down'. And if I press you regarding the undermining 'might' counterfactual 'if I were to jump, I might not come down', you agree that it has opposite truth value (we will have you disagree with this soon enough): in normal conversational contexts it is false. If I then press you with quantum mechanical and other anomalous possibilities, you agree that the 'might' counterfactual becomes true, *but only relative to a changed context*. So the truth-value of 'if I were to jump, I would come down' is context-sensitive. But since most contexts are normal ones, we can happily say that in most contexts it is true. Or so you may say.

You may respond to my argument from disjunctive antecedents in a similar way. 'If I were to jump, I would come down' is true, you say. But the equivalent counterfactual with the long antecedent that spells out all the various ways of my realizing a jump changes the context by making salient possibilities that we previously ignored. And in that new context, both counterfactuals are false.

In another version of this ‘counterfactuals are context dependent’ objection, you may allow both the ‘would’ and the ‘might’ counterfactuals to go *indeterminate* in some contexts, as well as true in others and false in still others. Still, you stick to your previous guns that in normal conversational contexts, ‘if I were to release the cup, it might not fall’ is false—and so on.

I reply: The oft-heard slogan that ‘counterfactuals are context-dependent’ glides nicely off the tongue, the way that slogans often do. Moreover, it’s meant to have consequences for our various practices—for example, van Fraassen (1980) thinks that it has consequences for the place of counterfactuals, or otherwise, in science. But what does the slogan mean exactly? Is it that *the truth values of all* counterfactuals are context-dependent? That’s clearly false: consider any counterfactual of the form p □→ p, and more generally any of the counterfactuals whose context-*independent* truth I will concede in §6. Is it that *some* counterfactuals have context-dependent truth value? That’s obvious: “if I were you, I’d lose that tie” depends on the contextually-sensitive terms ‘I’, ‘you’, and ‘that tie’.

But let’s set aside the context-dependence of counterfactuals that is parasitic on context-dependence of something else (indexicals, gradable adjectives, quantifier domains, epistemic modals, knowledge ascriptions, or what have you). To be of interest, and of potential trouble for me, the thesis must be something like “*most* counterfactuals have context-dependent truth value”—perhaps, most counterfactuals on my imagined transcript of all counterfactuals uttered in human history are context-dependent? I haven’t seen a careful argument for that. I suppose the usual argument, to the extent that there is one, is roughly this: Quine’s famous ‘Caesar’ example is meant to be typical: “If Caesar had invaded Korea, he would have used catapults/nuclear weapons”—context-dependent! And this counterfactual is supposedly context-dependent because in one context we may focus on Caesar’s belligerent tendency to use the most powerful weapons at his disposal (which may include nukes), and in another context we may focus on the historical facts about his weapons (which exclude nukes).

I have several replies to this argument. Firstly, it is not clear that examples like the Caesar counterfactuals *are* typical. The even more famous "If Oswald had not shot Kennedy, then somebody else would have" is apparently straightforwardly *false*, independent of context.[18] Counterfactuals like this seem no less typical to me. Likewise, my considerably less famous 'If the coin were tossed, it would land heads' surely has a context-independent truth-value—false, as I keep urging. Or again, 'If I were at least 7 ft tall, I would be exactly 7 ft tall' seems false, irrespective of context. So I'm not yet convinced that Quine's example is so typical. To be sure, a single example suffices to show that a complete semantic analysis for counterfactuals will need a contextual parameter, and that example may as well be Quine's; it's just that the parameter may be idle for many counterfactuals. But then the context-dependence of counterfactuals does not yet pose a threat to my conclusion that most counterfactuals are false; there needs to be a further premise about how often the parameter is activated.

Secondly, I agree that our similarity judgments about worlds may often be context-dependent, since what we hold fixed in those judgments may be context-dependent. But I have questioned similarity-based accounts of counterfactuals.

[18] If you are a conspiracy theorist about the Kennedy assassination, append the words "at exactly the same instant", and you should agree that the resulting counterfactual is false, independent of context.

Thirdly, it does not follow from the context-dependence of similarity judgments that the *truth value* of the ‘Caesar’ counterfactual is context-dependent—for it may well be that it comes out false *whichever* way we contextually resolve our similarity judgments. Indeed, that is exactly what we should expect, given my arguments. Focus on Caesar’s belligerent tendency if you like. Still, there would have been a chance of him *not* using nukes if he had invaded Korea; still, he *might not* have used nukes if he had invaded Korea (he might have used chemical weapons instead). Focus on the historical facts if you like. Still, there would have been a chance of him *not* using catapults if he had invaded Korea; still, he *might not* have used catapults if he had invaded Korea (he might have used slingshots instead). So the corresponding ‘would’ counterfactuals come out false either way—false in *different* ways, but false nevertheless. So much for context-dependence!

More generally, even granting that our similarity judgments about worlds may be context-dependent, it does not follow that the truth-values of the everyday counterfactuals that we utter are context-dependent. They may well be stable across context shifts—and my position predicts that most of them are stably false. Compare: while our judgments of ‘tallness’ are context dependent, the truth value of ‘all tall people have two heads’ is stable across context shifts—stably false.

And what of the alleged context-dependence of the *truth value* of ‘if I were to jump, I would come down’—true in ordinary contexts, you say, and false (only) in extraordinary contexts in which quantum mechanical or other anomalous possibilities are salient? Counterfactuals just don’t seem to behave like paradigmatic examples of context-sensitive parts of language, like pronouns, gradable adjectives, quantifier domains, and so on. Let’s take the case of quantifier domains, since arguably it provides the best model of how the context-dependence of counterfactuals is supposed to work: for a given counterfactual we somehow restrict the set of antecedent-worlds under consideration, and the restriction is supposed to be context dependent. (For example, quantum tunneling worlds are excluded in ordinary contexts, included in contexts where such tunneling has been made salient.) But I believe my remarks generalize to all context-dependent language.

When I say at the start of a party ‘All the beer is in the fridge’, there is clearly a tacit restriction on my domain of quantification. When you say, “Whoa! Big fridge!”, we both know that you are joking: I meant that *all the contextually salient beer* is in the fridge—as it might be, all the beer that we bought this afternoon—not all the beer in the world. If need be, I can make the contextual parameter, the domain restriction, explicit; I can give a context-*insensitive* paraphrase of what I said. A latecomer to our conversation may not know which beer I was talking about. I can clarify: “all the beer that we bought this afternoon”. Or if you and I are talking at cross purposes—you thought I meant all the beer that the guests brought—again I can clarify. Later in our drunken party, conversation may turn to a free-ranging discussion of all the beer in the world. The context has changed. I can no longer truly say: “All the beer is in the fridge”—that really *would* deserve a “Whoa!” But recalling what I said at the start of the party, I can rightly maintain: “What I said then was true.” I am under no pressure to *retract* it. And I don’t have any sense of *enlightenment* once we have entered the free-ranging context; I don’t regard it as in any sense a *superior* or *more thoughtful* context.

None of these phenomena carry over to counterfactuals. I say at the start of the party: “If I were to jump, I would come down”. It is quite unclear what would be a context-insensitive paraphrase of that. You might try: “If I were to jump *in a normal way*, I would come down”, but “normal” is still context-sensitive (it’s a gradable adjective), and it is at best shorthand for an open-ended list (no quantum tunneling, no Maxwell demon monkey-business, no sudden gusts of wind) that I could never make explicit. “Normal” could similarly be in need of clarification for a latecomer unfamiliar to our context, or for you if we are talking at cross purposes, and my attempts at clarification will always be incomplete. Later in the party, our conversation turns to quantum tunneling (the lives of the party that we are!). I cannot truly say: “If I were to jump, I would come down”. But nor can I can rightly maintain of what I said at the start of the party: “What I said then was true.” I *am* under pressure to retract it. And I do so because I feel comparatively enlightened in my new, more demanding context: it is a superior, more thoughtful context, because it pays due recognition to the chanciness of my hypothetical jump, in virtue of its appreciation of the physics that bears on it, *and that always did* bear on it, whatever our context.

Reflect on why you think the counterfactual is false in the latter contexts. A good reason, I submit, is that the *chanciness* of my coming down is made salient in these contexts, and it undermines the claim that I *would* come down, if I were to jump. But the chanciness was there all along, before we made it salient or attended to it. Compare: Suppose a friend is poised all along to try to spare me a fall if I were to jump; she would probably fail, but there is some chance that she would succeed. Then it is immaterial whether you happen to be attending to her or not—irrespective of whether this possibility is salient or not to you, and whether you realize it or not, it is simply false that if I were to jump, I would come down. The chancy quantum mechanical and other anomalous possibilities are other ways the fall might fail to occur, whether you attend to them or not. But if you think that chanciness does *not* undermine the counterfactual, then why do you feel any inclination to retract the counterfactual when we move to a context in which the chanciness is salient? Either a given counterfactual is compatible with the chanciness of its consequent, or it is not. If it is, then you are wrong to retract it after attending to that chanciness. If it is not, as I claim it is not, then you are wrong to endorse it before attending to that chanciness.

But suppose for the sake of the argument that a broad class of counterfactuals really have context-dependent truth-values. (It had better be broad to threaten my thesis.) We might follow Lewis's account of knowledge (1996), for example, in contending that context determines which possibilities are and are not properly ignored. In a typical conversation about cup dropping, we may properly ignore bizarre quantum mechanical possibilities; in a typical conversation about jumping, we may properly ignore bizarre statistical mechanical possibilities. But once I draw your attention to them, you cannot ignore them (for a while, anyway), let alone properly ignore them. And as long as they are live, they will underwrite various outlandish ‘might not’ counterfactuals that oppose corresponding mundane ‘would’ counterfactuals. This has the unintuitive consequence that a good way to make your counterfactual assertions come out true is to make sure that you and your interlocutors are inattentive, ignorant, or unimaginative, and the more so, the better.[19]

But suppose I’m wrong about all of this, and that counterfactuals really have context-dependent truth values. Then there is still a striking result: most counterfactuals are easily *made* false: as easy as it is to shift to a context in which bizarre possibilities are salient. An extreme case of this, also reminiscent of Lewis on knowledge, is that philosophical contexts will be especially liable to make counterfactuals false, for in such contexts bizarre possibilities are fair game—bad news for those philosophers *in* such contexts who want to use counterfactuals to analyze other concepts. We will return to this point at the very end. And never mind conceptual analysis; we had better be careful, when asserting that counterfactuals are *entailed* by other things (e.g. the laws of nature), that we either embrace the context-dependence of these entailers, or else explain how their context-independence fails to transmit to their entailments.

Suppose that the truth-maker for some claim we hold dear is not merely a counterfactual in isolation, but a complex pattern of counterfactuals. The contextualist had better hope that they all come out true in the relevant context. If they are too unstable, too sensitive to context, and the context changes sufficiently from one to another, then the dearly held claim may be undermined. For example, suppose that causation involves chains of counterfactual dependences (à la Lewis 1973), and consider a case in which the chain is long: E is counterfactually dependent on C_n, which is counterfactually dependent on C_{n-1}, which is …,

[19] Compare Elgin (1988).

which is counterfactually dependent on C_2, which is counterfactually dependent on a salient cause C_1. If the individual counterfactuals are too fickle, there may be no single context (let alone the relevant context) in which they all come out true. Then there is no single context (let alone the relevant context) in which the claim that C_1 causes E comes out true. This will have ramifications in turn for concepts that depend in turn on causation—say, moral responsibility. Indeed, the problem is writ large when a concept dear to us involves an entire network of causal chains—think of mental states on a functionalist analysis. The point generalizes beyond these illustrations: one should not be glib about the putative context-dependence of counterfactuals while being sanguine that complex patterns of counterfactuals, across which contexts may vary significantly, underwrite important claims of ours.

And even if we grant that most counterfactuals have context-dependent truth values, that it still quite compatible with my thesis that most counterfactuals are false. Recall my imagined transcript of all counterfactuals ever uttered in human history. It is quite compatible with contextualism that most of them were *uttered in contexts in which they were false*. Indeed, it is compatible with contextualism that all counterfactuals, as a matter of fact, are false in their context of utterance.

At this point the contextualist may want to invoke some version of a principle of charity—roughly, a constraint on interpreting someone else's utterances so that most of them come out true, and where contextual factors make a difference, they are resolved with a presumption in favour of the truth of what is said. As a semantic thesis, I find the principle of charity too good to be true, and I don't know a charitable way of understanding it that will save it. It surely diminishes the achievement of getting things right. Think of how hard this can be in daily life, let alone in science. We know, for example, that Newtonian mechanics is a *false* physical theory; it would be absurd to suggest that we just need to interpret it charitably in order to see that it is true after all! We should not be credited with too many true utterances, and we should get more credit for the true utterances that we do make. The principle of charity also risks making disagreement harder to come by than it should be—indeed, it may be *uncharitable* to regard context shifts as ensuring that apparently disagreeing parties are each speaking truly, especially when they take their disagreement to be genuine.

As a pragmatic strategy or heuristic for felicitous interactions with one another, something like the principle of charity has more going for it, although then it poses no threat to my semantic thesis about the counterfactuals' truth values. And even then the principle seems a little off the mark. More plausible is something more like the principle of humanity—roughly, when interpreting someone else's utterances, we should regard her beliefs and desires as connected to each other and to reality in a way that makes her as similar to ourselves as possible. This makes more room for false beliefs in the face of misleading or incomplete evidence, as is all too common.

In any case, I am not convinced that attributing falsehoods to each other need be uncharitable or inhumane. Falsehood gets a bad rap! Falsehoods may have various desirable features—they may be justified, useful, acceptable, assentable, assertable, and so on. Listen carefully to a typical conversation. You should be struck by how much of what is said is literally false, but understandably so because of other purposes that are served—think of exaggerations, jokes, short-cuts, approximations, imperfect paraphrasing, irony, metaphor, diplomacy, and so on. False propositions may play essential roles as premises in knowledge gained by inference. (See Warfield 2005, Fitelson 2010.) Arguably, even much of science is literally false, but close enough to true about enough of the propositions that we care about, to do the job we require of it. There should hardly be some special presumption in favour of the truth of counterfactuals that we utter, given so much falsehood in our utterances across the board! And we seem to be especially bad at modal reasoning—witness the literature pioneered by Kahneman and Tversky on how bad people are at probabilistic reasoning. Why think that we're especially *good* at counterfactual reasoning?

Still, I do think that counterfactuals fare especially poorly even compared to our generally low strike rate of truths in what we say, even the modally charged things we say. That should hardly come as a surprise, given my arguments. After all, the pronouncements of quantum mechanics and of statistical mechanics come as a shock. Startling quantum mechanical or entropy-decreasing possibilities fly in the face of folk wisdom, so we should hardly expect folk intuitions to be a reliable guide to the counterfactuals that they undermine. That said, in §8 I will explain how our false utterances of counterfactuals can be vindicated by closely related truths. In the meantime, I suggest that if most of our counterfactual utterances are false, they are in surprisingly good company.

Returning to the contextualist challenge to my position: even granting the context dependence of the truth-values of counterfactuals, I have as fall-back positions: *most counterfactuals are easily made false*, and they are *especially easily made false in philosophical contexts*, *important claims of ours that are underwritten by complex patterns of counterfactuals risk being false*, and *most counterfactuals may be uttered in contexts in which they are false*. These positions are still somewhat unsettling.

5.2 The counterfactuals are indeterminate/lack truth value rather than being false

You may say that in indeterministic cases such as I have discussed, and in deterministic cases with imprecise antecedents, there is no determinate fact of the matter of what *would* happen. You may insist, then, that these counterfactuals are not false, but *indeterminate*. Or you may say that such counterfactuals simply lack truth value. For example, starting with the coin that will never be tossed, 'Toss □→ Heads' is neither true nor false, where I too quickly concluded that it was false. And when I took you down the slippery slope from there, through biased coins and lotteries to billiard balls, jumps and cups, at each point you may judge the counterfactuals to be neither true nor false. Or so you say.

I reply: This still yields a striking conclusion, namely, most counterfactuals are *not true*. Striking, because it still undermines much of what we say, apparently *taking* such counterfactuals to be true. Try telling people on the street that 'If I were to jump, I would come down' is *not true*, and see what looks you get!

[I NEED TO INSERT HERE A DISCUSSION OF PROPONENTS OF THE VIEW THAT THE COUNTERFACTUALS HAVE **INDETERMINATE** TRUTH VALUE]

As I argued in §2.2.1, the incompatibility of 'would' and 'might not' counterfactuals suggests that the 'woulds' are not indeterminate, but downright false. For example, 'Toss ◊→ ¬Heads' is not merely indeterminate, but *true*. Arguably, this entails that 'Toss □→ Heads' is *false*. Let me add now that if 'would' and 'might' counterfactuals are duals, then the argument for this is especially straightforward. By that duality, 'Toss □→ Heads' is equivalent to '¬(Toss ◊→ ¬Heads)'. Suppose for reductio that the former is indeterminate. Then so is the latter. And if the latter is indeterminate, then so is its negation, 'Toss ◊→ ¬Heads'. But that might-counterfactual is *not* indeterminate, but *true*. I subscribe to the would/might duality, and I am in good company, so I stand by my claim that most counterfactuals are *false*.

The duality of 'would' and 'might' counterfactuals similarly casts doubt on no-truth-value accounts of counterfactuals. (Since the no-truth-value theorist may not be happy with our speaking of the equivalence of p □→ q and ¬(p ◊→ q), we may speak instead of them necessarily have the same probability.) Simply replace "indeterminate" by "lacking in truth value" in the argument that I have just given.

Edgington (2008) is a no-truth-value theorist about counterfactuals—she thinks that like indicatives, they are governed by a version of the Adams Thesis that the probability of a conditional is the corresponding conditional probability: the probability of p □→ q is $P(q \mid p)$. But one wonders what the probability of a counterfactual *is,* if not its probability of truth. And this 'no truth value' account is surely not true across the board. What about counterfactuals of the form p □→ p, and more generally, those whose antecedents entail their consequents? In §6 I will concede that various kinds of counterfactuals are true. So Edgington's thesis cannot be anything stronger than *most counterfactuals lack truth value*. And what about a counterfactual with a true antecedent and false consequent? It is presumably *false*, not merely improbable. To be sure, one might question whether it is really a *counterfactual* at all, since its antecedent is not contrary-to-the fact. If it is not, then this will only help my own cause in §6.5, when counterfactuals with true antecedents and true consequents appear to pose a threat to my position. If they are not really counterfactuals in the first place, then I need not worry about them. But if they are, then we surely have more examples of counterfactuals with genuine truth values, contra no-truth-value theorists.

There's also the familiar Frege-Geach problem of how counterfactuals are meant to embed in various contexts, such as in Boolean combinations, or modal contexts, or how they iterate. Sometimes iteration can be explicit, sometimes implicit. Take your favourite concept that you think should be analyzed in terms of counterfactuals—for definiteness, let it be causation. Then the superficially uniterated counterfactual 'if it were that p, it would be that C

causes *E*’ unpacks as an iterated counterfactual. Furthermore, counterfactuals can appear in arguments that offhand seem to be valid or invalid, and I don’t just mean some kind of weaker ‘probabilistic validity’, à la Adams. And if no-truth-value theorists want to analyze other things in terms of counterfactuals, as philosophers so often do, what will they say about those other things—causation, knowledge, perception, personal identity, laws of nature, dispositions, and what have you? That claims involving these things don’t have truth value either? I’ll return to this point at the very end.

But if I’m wrong about all this, the fall-back position that most counterfactuals are *not true* is disturbing enough.

5.3 The ‘might not’/’would’ clash is merely pragmatic

My argument for the falsehood of most counterfactuals in §2.2 began with the collision between a ‘would’ and an opposed ‘might not’. I asked you to try saying:

> “If I were to toss the coin, it MIGHT land Tails, and it WOULD land Heads if I were to toss the coin” (*)

and to consider whether you’ve said something consistent. You may agree that (*) sounds odd, but deny that the oddness is logical. It may instead be some kind of *pragmatic*, rather than semantic, inconsistency—the oddness that apparently underlies Moore paradoxical sentences such as ‘It’s raining and I don’t believe it’s raining’. In that case, both ‘Toss ◊→ Tails’ and ‘Toss □→ Heads’ could both be true. Then I lose my argument from the truth of the second conjunct to the falsehood of the first. And so it goes for the other counterfactuals that I went on to discuss. For example, ‘Jump □→ Fall’ and ‘Jump ◊→ ¬Fall’ are semantically consistent; they are merely pragmatically inconsistent. Or so you may say.

More specifically, DeRose (1999), following Stalnaker (19xx), argues that 'if it were the case that p, it might be that q', should be analyzed as:

$\Diamond_e(p \Box\!\!\rightarrow q)$

where '$\Diamond_e$' symbolizes 'it is epistemically possible that'. Notice that this renders the 'might not' counterfactual consistent with the corresponding 'would' counterfactual. DeRose then appeals to the pragmatics of assertion, in particular the Williamsonian thought that "you represent yourself as knowing a fact if you flat-out assert it" (388). Suppose you assert 'if it were the case that p, it would be that q'. Then you pragmatically convey to your audience that know $p \Box\!\!\rightarrow q$, and hence that $\neg(p \Box\!\!\rightarrow q)$ is not an epistemic possibility for you—that is, you convey $\neg\Diamond_e\neg(p \Box\!\!\rightarrow q)$. But you say something semantically inconsistent with this if you then go on to assert 'if it were the case that p, it might be that $\neg q$', that is, $\Diamond_e(p \Box\!\!\rightarrow \neg q)$. For this entails $\Diamond_e\neg(p \Box\!\!\rightarrow q)$, the negation of what you conveyed. We thus have an ingenious explanation of inescapable clashes. They involve the same kind of pragmatic inconsistency as we find in assertions of the form 'X, and it might be that not-X'—semantically consistent, but representing oneself in a way that is inconsistent.

The Stalnaker/DeRose account is by no means the only possible pragmatic account of the might-not/would clash, but it is particularly carefully worked out, and it will serve as something of a case study.

I reply:

I will begin with some worries about pragmatic approaches in general, turning then to some specific objections to Stalnaker/de Rose analysis.

Presumably, we would like to explain inescapable clashes even when they are not flat-out asserted—for example, when they are merely thought, or supposed, or mentioned, or reported, or the consequents of iterated conditionals, or are items in a data base that we are trying to assimilate. We would also like to explain them when the ‘might not’ counterfactual is explicitly *metaphysical,* rather than epistemic. In the cases I have imagined, that reading is available, and arguably even primary. The inferences from instances of indeterminism or of indeterminacy to metaphysical ‘might nots’ strike me as secure, and it is unclear to me what bearing pragmatic considerations could have on them. And however the pragmatic inconsistency of the clashes is cashed out, the fact that they are *clashes* that are pragmatically *inconsistent* tells us that there is still something seriously defective with asserting both the ‘might not’ counterfactuals and the corresponding ‘woulds’. Yet I claim that there is nothing wrong with asserting the ‘might nots’; so still, the ‘woulds’ must take the fall—now, the fall of unassertability.

Turning now to the specifics of the de Rose/Stalnaker analysis: For starters, it does not do justice to the surface grammar of the ‘might counterfactuals’, in which the ‘might’ has narrow scope. To be sure, this counts for little; still, it’s a point against, rather than a point for, the analysis. More importantly, it founders on our cases of indeterminism and indeterminacy. ‘If the coin were tossed, it *might* land Tails’ does not say the same thing as ‘It is epistemically possible that: if the coin were tossed, it *would* land Tails’. For the former is true, and the latter is surely false according to anyone who knows that the coin toss would be an indeterministic process. The ‘would’ claim is incompatible with their knowledge of the chanciness of the coin. It is the very nature of chancy processes that there is no fact of the matter of how they *would* eventuate. Stalnaker’s own story of Tweedledee and Tweedledum shows us that.

Similarly, ‘If I were at least 7 ft tall, I *might* be 7 ft 1 inch tall’ can reasonably be regarded as true, while ‘It is epistemically possible that: If I were at least 7 ft tall, I *would* be 7 ft 1 inch tall’ can reasonably be regarded as false. Given the indeterminacy regarding my hypothetical height, I may know that 7 ft 1 is a live option; but I know that it is *not* the only live option, so I know that the ‘would’ counterfactual is too committal to be true.

But let’s suppose for the sake of the argument that DeRose’s analysis of ‘might’ counterfactuals is correct. This does not yet show that the counterfactuals that we ordinarily take to be true are in fact so. For it does not follow from the fact that a given ‘might not’ counterfactual is true, and that the corresponding ‘would’ counterfactual is compatible with it (despite clashing with it pragmatically), that the ‘would’ counterfactual is true. (Compare: it is true that grass is green, and compatible with this that snow is black; it does not follow that it is true that snow is black.) We may grant that there is a possible world in which both the ‘might not’ and ‘would’ counterfactuals are true, without granting that the *actual* world is such a world. At best, DeRose has shown that our pre-theoretical belief that various ‘would’ counterfactuals are true is *tenable*, not that it is *correct*. And I have given other arguments that this pre-theoretical belief is *incorrect*.

What about DeRose’s explanation of the pragmatic inconsistency? Arguably, one represents oneself merely as *believing* something when one asserts it, rather than knowing it. When one really does want to represent oneself as knowing something, explicitly saying so is not redundant: "The pub closes at midnight. Trust me: I *know* that it does." And it's not the difference of now representing oneself as knowing that one knows—that's too fancy for the folk, who often speak this way (even outside pubs).

Finally, and most importantly, the DeRose proposal is still open to a kind of counterfactual skepticism, one that I contend is rather more puzzling than the one that I am defending. Whenever a 'might not' counterfactual is true, the corresponding 'would' counterfactual cannot be *known,* even if it is in fact true. (Cf. Eagle 2007.) But then the 'would' counterfactual cannot reasonably be asserted either, for in doing so one would *misrepresent* oneself as knowing it. So we quickly reach the startling conclusion that *most counterfactuals are unassertable*. But this flies in the face of our linguistic practices, which are a surer guide to assertability than they are to truth. My skepticism about the truth of counterfactuals seems tame by comparison! By my lights, DeRose has things exactly back to front: while I regard most counterfactuals that we utter to be false but assertable (more on that in §8), DeRose must regard most of them as unassertable, even though he thinks that they are true. He is throwing the baby out with the bath water.

5.4 'Would' and 'might-not' counterfactuals are semantically and pragmatically consistent

You may insist that there is no inconsistency whatsoever between the 'would' counterfactuals and the corresponding 'might not' counterfactuals—neither semantic, nor pragmatic. Indeed, in my central examples, both the 'woulds' and the corresponding 'might nots' are true. For instance, 'if I were to jump, I would come down' is true (as commonsense would have it), *and* 'if I were to jump, I might not come down' is also true (as physics would have it). Or so you say.

I reply: I hope that you will still agree that there is *some* tension between these sentences. If not, our linguistic intuitions are at such odds that I'm not sure how we could move beyond an impasse. And if we do agree at least that much, then I offer you the challenge of doing justice to the tension that even you hear in (*), and other examples of inescapable clashes. For instance, if you think that p □→q involves the nearest p-worlds, whereas p ◊→¬q involves casting our net further afield to include more distant p-worlds, as Heller (19xx) does [CHECK!!] then it is unclear why there should be any tension *at all* between them.

But if there is really is no tension at all between them, their standing alongside each other in perfect harmony, then we still reach an interesting result, my final safety net. In that case, we still have an argument against the Lewisian (1973) semantics for counterfactuals, according to which the 'would' and 'might not' counterfactuals are in such tension that they are contradictories. And we have an argument against any proposal for there being a pragmatic inconsistency between assertions of them, such as that of Stalnaker/DeRose. These are not my arguments, but they may still be worthy of attention.

*　　　*　　　*　　　*　　　*　　　*

Much philosophy is an attempted demolition of commonsense followed by damage control. I think that it is an item of commonsense that various ordinary counterfactuals, such as 'if I were to jump, I would come down' are true. I have argued that commonsense is mistaken about them; indeed, they are not even indeterminate. So much for commonsense. Now it is time for some:

6. Damage control

There are limits to how startling my conclusion can legitimately be. I now concede that *some* counterfactuals are true; but I will maintain that this will not amount to all that much of a concession.

Let us begin with counterfactuals that are not just true, but necessarily so, in virtue of logical, mathematical, analytic, metaphysical, or nomological truths.

6.1 Strict conditionals

6.1.1 Necessary consequents

Counterfactuals with necessarily true consequents are trivially true (even if recognition of their truth may be non-trivial), where the 'necessity' at issue may be logical, analytic, mathematical, metaphysical, or nomological. Thus, I happily concede that the following counterfactuals are all true:

if the coin were to be tossed, it would be self-identical;

if the coin were to be tossed, all bachelors would be unmarried;

if the coin were to be tossed, nothing would be red and green all over;

if the coin were to be tossed, Fermat's last theorem would be true;

if the coin were to be tossed, water would be H_20;

if the coin were to be tossed, Hesperus would be Phosphorus;

if the coin were to be tossed, then nothing would travel faster than light;

and so on.

Counterfactuals whose consequents assert something true about the past may well go the same way:

if the coin were to be tossed (now), World War II would still have occurred,

and so on. They may well be just further instances of necessary consequents (given the actual historical facts): most philosophers agree that it is impossible to change the past.

6.1.2 Impossible antecedents

Likewise, I am prepared to concede happily enough that counterfactuals with necessarily false antecedents are trivially true:

if the coin were not self-identical, then it would land heads;

if the coin were not self-identical, then it would not land heads;

and so on.

Now, I am not sure that I am forced to make this concession. Counterfactuals with necessarily false antecedents may be used non-trivially in *reductio* reasoning, and this surely requires that they are not all vacuously true. Furthermore, consider counter-metaphysicals: ‘If there had been exactly 17 possible worlds, then Lewis’s views on possible worlds would have been correct in every detail.’ That hardly sounds true, since Lewis devoted much ink to arguing that there are infinitely many possible worlds. Or counter-nomologicals: ‘If gravity had obeyed an inverse cube law, then the planets would still have followed elliptical orbits’ is surely false.

But let us set these aside, for they will not affect my conclusion: if even such counterfactuals can be false, all the better for my case.

6.1.3 Necessary connections between antecedents and consequents

Likewise, I happily concede that counterfactuals whose antecedents necessitate their consequents—either in virtue of logic, or analyticity, or metaphysical necessity, or mathematics, or nomological truth—are trivially true:

if I were in Australia, then I would be in Australia or Morocco;

if were not married, I would not have a wife;

if the moon were made of green cheese, the moon would not be red all over.

if I were 7 ft tall, I would be $\sqrt{49}$ ft tall;

if Renée were drinking water, then Renée would be drinking H_20;

and so on.

Each such counterfactual is of the form $p \mathbin{\Box\!\!\rightarrow} q$, where $\Box(p \supset q)$ is true—the 'box' being an appropriate form of necessity. In other words, strict conditionals underwrite these counterfactuals. And while counterfactuals are sometimes called 'strong' conditionals (see e.g. Lewis 1980), strict conditionals are stronger still.

Really, *all* necessarily true counterfactuals are species of those whose antecedents necessitate their consequents, so we may regard this category as subsuming the others. For if $\Box q$, then for any p, $\Box(p \supset q)$. And if $\neg\Diamond(p)$, then for any q, $\neg\Diamond(p \mathbin{\&} \neg q)$, which is equivalent to $\neg\Diamond\neg(\neg p \text{ v } q)$, and thus to $\Box(p \supset q)$. (This in turn provides an argument for regarding counterfactuals with impossible antecedents as true after all: $\Box(p \supset q)$ is stronger than $p \mathbin{\Box\!\!\rightarrow} q$, and since the former is true when p is impossible, so is the latter. So I have now given arguments for both verdicts regarding such counterfactuals.) We can then characterize *all* of the true counterfactuals that we have so far identified as ones in which *the antecedent necessitates the consequent*. It remains to be seen whether there will be any other true counterfactuals to add to the list.

It should come as no surprise that all these trivially true counterfactuals resist my argument. For clearly the might-counterfactuals that are contrary to them are false. 'Might' counterfactuals with impossible consequents are trivially false:

if I were to toss the coin, it might not be self-identical,

and so on.

It is less obvious that ‘might’ counterfactuals with impossible antecedents are trivially false:

if the coin were not self-identical, then it might land heads,

and so on. For it is a little odd that the ‘would’ counterfactual with the same antecedent and consequent is true, while the seemingly weaker ‘might’ counterfactual is false. Recall that I was unsure about my concession to the truth (let alone necessary truth) of ‘would’ counterfactuals with impossible antecedents. So if ‘might’-counterfactuals like this one come out to be true, all the better for me.

Again obviously, might-counterfactuals involving the failure of a necessary connection are trivially false:

if I were in Australia, then I might not be in Australia or Morocco,

and so on. In general, if $\Box(p \supset q)$ is true, then $\Diamond(p \,\&\, \neg q)$ is false, so $p \Diamond\!\!\rightarrow \neg q$ is false. If there is no world in which p and $\neg q$ are both true, then still less is there a $p \,\&\, \neg q$ world at least as close as any $p \,\&\, q$ world, as the Lewis semantics for $p \Diamond\!\!\rightarrow \neg q$ would have it.

Sometimes, the truth of a necessarily true counterfactual is secured by a contingent truth about the world. For example: in fact, the coin landed heads. I say, truly: “If I had bet on heads, with the coin landing as it in fact did, then I would have won.” But what underwrites this counterfactual is one in which the necessitation is laid bare: “If I had bet on heads, with the coin landing heads, then I would have won.” And I will happily concede its truth, too.

My concessions are happy, because they are not counterfactuals that you hear uttered on the street, or indeed outside a philosophical discussion (not that you would hear them much *inside* a philosophical discussion, either). They will appear relatively rarely on my imagined transcript of all counterfactuals ever uttered or written in human history.

6.2 Counterfactuals under determinism with sufficiently precise antecedents

Let us explore further an interesting subclass of counterfactuals whose truth is secured by a nomological connection between antecedent and consequent.

I argued earlier that not even determinism suffices to save counterfactuals from falsehood: the problem was that vaguely specified antecedents encompass initial conditions that yield anomalous results. What is needed is determinism *plus* sufficient precision in the antecedents regarding the initial conditions, relative to the consequents. That way, the antecedents necessitate the corresponding consequents, and we have more instances of type 6.1.3.

A counterfactual about my jump will be true provided that the antecedent together with the (deterministic) laws of nature imply the consequent. Thus, a counterfactual whose antecedent fully specifies my jump, molecule-by-molecule, will be true provided that all nomologically possible trajectories consistent with that specification result in the truth of the consequent—as it might be, ‘I fall’. A counterfactual that specifies my jump more imprecisely may still be true, but only if there is sufficient imprecision in the consequent to tolerate it. As we saw above, even the seemingly imprecise consequent ‘I fall’ was too precise to be nomologically implied by the antecedent ‘I jump’, so the corresponding counterfactual was false. To make it true, either we have to precisify the antecedent so as to rule out all anomalous initial conditions that result in my not falling, or we have to *imprecisify* the consequent so as to be compatible with all of them. Precisifying the antecedent sufficiently will be quite a task; while it may not require a molecule-by-molecule specification of the jump, it will require a lot of fancy footwork nonetheless, presumably intractably so. Imprecisifying the consequent sufficiently is easier but risks rendering the counterfactual pointless—e.g., “if I were to jump, something would happen to me”. Of course, we could secure the truth of the counterfactual by building in a logical connection between antecedent and consequent—e.g., “if I were to jump according to initial conditions that result in my falling, then I would fall”. But now the counterfactual is trivial.

So our options here for producing true counterfactuals—precisifiying the antecedent, imprecisifying the consequent, or building in a necessary connection between antecedent and consequent—yield counterfactuals that are respectively intractably complicated, pointless, or trivial. In any case, they will hardly appear on my imagined transcript.

6.3 Counterfactuals with probabilistic consequents

There are many other counterfactuals that *may* be non-trivially true, namely those that explicitly state the appropriate probability in the consequent itself. ‘If the coin were tossed, it would land heads with chance 1/2’ may well be true for all I’ve said. There may well be a tiny real number, $\varepsilon > 0$, such that ‘if I were to jump in the air, I would come down with chance $1 - \varepsilon$’ is true. And so on.

There is no contradiction between a counterfactual with a probabilistic consequent and the opposite ‘might’ counterfactuals, even when the probability is high, just as there is no contradiction between ‘$P(X) = x$’ and ‘*not* X’, even when the probability is high. There are precious few valid inference rules taking us from modal claims to probabilistic claims; there are even fewer—namely, none—taking us from probabilistic claims to claims about how things actually turn out. The only thing that can contradict a probability statement is a contradiction, or another probability statement (that attributes a different probability), or something that entails such a statement. This is true even if $x = 1$ or 0, as we saw in the example of the infinite sequence of coin tosses, all of which might land tails. So we have to countenance the possibility that various counterfactuals with probabilistic consequents are true.

I suspect that when such a counterfactual is true, it is a special case of a counterfactual with a nomological connection between antecedent and consequent—that is, once again a case of 6.1.3. The tossing of a coin is not lawfully connected to its outcome (under indeterminism); but it may well be lawfully connected to the *chance* of such an outcome.

So as before, perhaps I cannot appeal to a might-counterfactual that involves the failure of the lawful connection. Perhaps it is not true to say: 'if the coin were tossed, it might not land heads with chance 1/2', since the laws may determine the chance of heads to be 1/2. In that case, my argument from undermining might-conditionals will not go through. Similarly for all of the counterfactuals whose consequents are chancy, discussed in §2, as long as the correct chance is stated in the consequent. So counterfactuals with chancy consequents may be non-trivially true.

Then again, they too may be false, and for five different reasons.

i. The chance value claimed may be incorrect

The most obvious reason is that there is only one *correct* value for a given chance, and uncountably many *incorrect* values. The truth of 'If the coin were tossed, it would land heads with chance 1/2' implies the falsehood of 'If the coin were tossed, it would land heads with chance *x*' for all $x \neq 1/2$, since the chance function is a *function*.

What we really mean is "If the coin were tossed, it would land heads with chance *roughly* half". Given that there is some chance that a real coin lands on its edge, we can be confident that the chance of heads is *not* ½. It would be remarkable if we got the chance *exactly* right, to infinitely many decimal places of accuracy. Indeed, if the true chance is transcendental, then with very few exceptions (involving π, e, and their kin) we cannot even express it with our current linguistic resources. So we cannot express any corresponding true counterfactual involving it, either.

ii. The antecedent may not be sufficiently precise

Secondly, the antecedent may specify the conditions too vaguely to yield a unique chance. Indeed, this is surely true of the coin-tossing example (so an unqualified concession three paragraphs ago to its truth would have been premature). If the coin were tossed very feebly an inch above the ground, it might not land heads with probability 1/2; if the mathemagician Persi Diaconis were to toss the coin, it might not land heads with probability 1/2… Chances must be relativized to a chance set-up; but if the set-up is incompletely described, there may be considerable leeway in the resulting chances.

This is closely related to the argument I gave in §4. There I considered counterfactuals under determinism, and contended that an insufficiently precise antecedent could be compatible with anomalous results inconsistent with the consequent of a given counterfactual. Reverting to counterfactuals with probabilistic consequents is a way of recovering determinism at one step removed, as it were—for while there is indeterminism governing the consequent, the connection *between* the antecedent and a *probabilistic* consequent is supposed to be perfectly deterministic. And much as imprecision afforded a way of undermining the truth of counterfactuals under determinism, so it affords a way of undermining the truth of counterfactuals with probabilistic consequents.

iii. The chances may be chancy

Thirdly, it could be that the chancy consequents are themselves chancy—that there are *higher-order* chances attaching to the propositions in question. Chances attach to propositions; a typical chance statement has the form

$$ch(X) = x.$$

(Relativize this to a time, if you like.) However, that chance statement itself picks out a proposition—call it *Y*. For all we know, *Y* is itself in the domain of the chance function:

$$ch[Y] = y,$$

that is,

$$ch[ch(X) = x] = y.$$

There is an issue here of just how profligate chances are—just how big that domain is. But supposing that there is no restriction on the set of propositions that can be the bearers of chance, then we have no obstacle to such self-referential statements of chance. Moreover, such higher-order chances may be non-trivial, intermediate in value. This would seem to be a live possibility on a 'Humean supervenient' account of chance—see Lewis (1994) and Hoefer (20xx)

Perhaps, then, a counterfactual with a second-order chance statement as its consequent is true—say, 'if I were to toss the coin, then with chance 0.99993 it would land heads with chance 1/2'? But the problem, if it is a problem, may arise again one level up. The 0.99993 figure may itself be chancy. And so on. The staircase of non-trivial higher-order chances may never end. Moreover, counterfactuals with probabilistic consequents are hardly ordinary; with each added step up the staircase, they become even less so. So even if there is hope that somewhere up the staircase we reach true counterfactuals, they will hardly swell significantly the ranks of those true counterfactuals that are ever uttered or written.

iv. There may be chance gaps

Fourthly, unless chances are maximally profligate, attaching to *all* propositions, there are *chance gaps*—propositions that don't receive any chance at all. We've already seen some candidates for such propositions: those picked out by chance statements. Elsewhere (2003) I have argued that there are other candidates. I focused especially on *free acts* and *non-measurable sets*.

It may well be that *free acts,* such as my raising my foot now, simply don't receive any chance value at all. After all, while chances should not be identified with (limiting) relative frequencies (see e.g. Hájek 1996), such frequencies are typically good evidence for the values of chances, and they typically nearly coincide. Yet we can easily drive the (limiting) relative frequencies of our free acts to whatever value we want (and in the limit, to no value at all), *in virtue of their very freedom*.

Similarly, certain symmetric experiments over uncountable spaces give rise, at least in principle, to *non-measurable sets:* sets that cannot be assigned any probability at all, consistent with certain natural looking assumptions. Randomly throwing an infinitely fine-tipped dart at a representation of the [0, 1] interval is arguably such an experiment. Such sets would be chance gaps. See Hájek (2003) for more explanation and defence.

The upshot is that various propositions may simply fail to receive a chance value at all, under a given counterfactual assumption. In that case, any counterfactual with that assumption as its antecedent, and *any* particular chance attribution to such a proposition as its consequent, is false. It claims that were that antecedent to be the case, the chance of the consequent would be such-and-such, when in fact that chance is undefined.

v. The chances may be imprecise

Finally, and now letting our two devices for undermining counterfactuals work in tandem, *chances* themselves might be *imprecise*. For it is plausible that chances are determined by the laws of nature, and if the laws themselves are imprecise, chances could inherit this imprecision. This would certainly seem to be a live possibility on a Mill-Ramsey-Lewis style account of laws as regularities that appear as theorems in a 'best' theory of the universe, as long as the criteria for what makes one theory better than another are themselves imprecise. (In Lewis' 1973 theory, for instance, the vagueness may enter in the standards for balancing the theoretical virtues of 'simplicity' and 'strength'.) Then nature may not determine a single best theory, but rather a multiplicity of such theories. Suppose, for example, that these equal-best theories disagree on the chance that a radium atom decays in 1500 years: for each real number *r* in the interval [1/3, 2/3], there is such a theory that says that the chance is *r*. Then the chance of this event may be imprecise over this interval.

There may be even more straightforward ways for the laws of nature to be vague. Perhaps some of the fundamental physical constants are not entirely precise—perhaps, for example, the gravitational constant is only fixed up to 100 decimal places. Or some of the fundamental physical properties might be vague. It would seem that the laws in which such constants or properties figure would then be rendered imprecise, and that chances, which are determined by the laws, are correspondingly imprecise. The chance that this coin lands heads on a given toss, for example, may not be a sharp number such as 1/2, but rather a set of such numbers, or perhaps an interval. In that case, even the counterfactual ‘If this coin were tossed, it would land heads with chance 1/2’ is false—it overreaches, claiming a sharp chance when there is none. The true counterfactuals concerning the coin have consequents that are imprecise: if I were to toss the coin, it would land heads with chance in such-and-such interval, or in such-and-such set, or with chance fixed to only-so-many decimal places. And such counterfactuals are certainly rarely uttered.

Indeed, they may need to be still more complicated than that in order to be true. Suppose we admit higher-order vagueness, so that not even the endpoints of the intervals, or the number of decimal places, are sharp. We could handle such cases by appealing to still more sophisticated devices, packing them explicitly into the consequent itself. Of course, that means that the counterfactuals are getting still less ordinary: where previously they had probabilistic consequents with sharp probabilities (which made them rarified enough), now the consequents have vague probabilities, suitably characterized by whatever devices for higher-order vagueness are needed. We have left street talk far behind.

*   *   *   *   *   *   *

In sum, even counterfactuals with chancy consequents can be false, since the chance specified can be incorrect, the antecedent might be too imprecise to yield a unique chance, there might be higher-order chances, there might be chance gaps, and there might be imprecise chances. So I may not even have to concede, happily or otherwise, the truth of various counterfactuals with chancy consequents. But I don’t mind much either way, because they will hardly appear on my imagined transcript. So we don’t have to settle these contentious points here.

We have seen earlier that all trivially true counterfactuals can be regarded as those whose antecedents necessitate their consequents. We can either precisify the antecedent, or tailor the consequent to the antecedent (getting the chance right), or imprecisify the consequent. We've been looking at technical devices for achieving these things. How do we do that while keeping the counterfactuals ordinary? But most counterfactuals that are ever uttered are ordinary.

We won't see the former device, precisifying, much in ordinary language, so let's go the other way. What imprecisifying devices do we have?

6.4 Comparative and qualitative counterfactuals

I *will* concede that there are counterfactuals with consequents that are more imprecise still than those I have considered that might count as ordinary—for example, some counterfactuals with *comparative* or *qualitative* consequents concerning probabilities. "If I were to jump, I would be much more likely to fall than not", "if I were to jump I would very likely fall", and so on may count as ordinary. If so, my concession is a tad less happy, but there it is. I did, after all, only claim that *most* counterfactuals are false.

But maybe even this concession is too hasty. I surmise that these counterfactuals would strike the ordinary person as rather odd. They would find puzzling their apparent coyness, their diffidence, in much the same way that they would find puzzling the remark, perhaps from someone who has read a bit of Hume, but not too much: "The sun will *probably* rise tomorrow". That may well be true, but our inner Gricean wants to hear asserted the stronger claim that we also believe to be true: "The sun *will* rise tomorrow. (Dammit!)" Likewise, that same Gricean wants to hear asserted "if I were to jump, I *would* fall. (Dammit!)", taking this (rightly) to be a stronger claim than its comparative or qualitative counterparts, and (wrongly) to be true. In short, I am not sure that ordinary people would find the latter counterfactuals so ordinary, and they would rarely utter them. So such counterfactuals that really are ordinary are false, and the corresponding ones that may well be true are not ordinary, and rarely uttered.

Still, eventually my resources for denying the truth of counterfactuals will run out, and some of the counterfactuals that remain are—I grudgingly concede—ordinary, and will make some appearances on the great transcript. Consider a counterfactual whose consequent's probability is easily and correctly characterized in comparative or qualitative terms, sufficiently informatively that even the Gricean in us is satisfied. The truth of "If Shakespeare had not written Hamlet, very probably nobody else would have" seems secure, as is its ordinariness. Indeed, there will be a spectrum of such cases, corresponding to the range of such qualitative or comparative probabilistic claims: "If Einstein had not come up with the theory of relativity, it is fairly likely that eventually someone else would have"; "if Gore had campaigned harder in Florida, he would have improved his chances of winning the election", and so on. Come to think of it, we have a way of capturing the whole spectrum at once: "if Kennedy had not been shot in Dallas, there would have been *at least some chance* of his serving his full term as president." But that sounds like just another way of saying the 'might'-counterfactual, "if Kennedy had not been shot in Dallas, he *might* have served his full term as president." (Yes, I know that events of chance zero can happen, and I even exploited their existence in §2. But they require infinitely many trials, something we do not have in the Kennedy example.) 'Might' counterfactuals, that is, correspond to 'would' counterfactuals with maximally imprecise probabilistic consequents. And of course I have already conceded the truth of various ‘might’ counterfactuals; in fact, doing so has been the linchpin of one of my central arguments!

6.5 Counterfactuals with true antecedents and consequents

It is a consequence of both the Stalnaker and the Lewis semantics for counterfactuals that ‘$X \mathrel{\Box\!\!\rightarrow} Y$’ is true whenever both X and Y are true. This may be plausible on a similarity-based semantics, in which ‘$X \mathrel{\Box\!\!\rightarrow} Y$’ is true iff Y is true at all X-worlds sufficiently similar to the actual world. For no world could be as similar to the actual world as itself—this assumption is sometimes called *centering*—and if X & Y is true at the actual world, we are apparently done. Then it would appear that I need to concede the truth of another large class of counterfactuals: those whose antecedents and consequents are true.

Well, maybe not. This quick argument from considerations of similarity of worlds is a little too quick. It is too quick by my lights, since I have questioned similarity-based accounts of counterfactuals. And it is too quick if the notion of ‘similarity’ at play is *not* simply that of commonsense, but rather some (quasi-) technical notion, as I discussed in §2.2.3. Moreover, it is debatable whether conditionals with true antecedents should even count as ‘counterfactuals’. After all, they are not *contrary-to-the-fact-uals*, whereas that’s what is supposedly distinctive about ‘counterfactuals’. This could quickly devolve into a rather boring terminological dispute. It’s not clear why this dispute should be resolved in the less favourable way for me, but let me simply agree to resolve it that way. That still won’t undermine my position.

So, does X & Y imply $X \mathrel{\Box\!\!\rightarrow} Y$, as centering would have it? Here are several reasons for thinking not.

Firstly, it already strains the ear to say that $X \Box\!\!\rightarrow Y$ is true when X and Y are true but are independent of each other. ‘If Canberra were the capital of Australia then the moon would have large craters’ is more likely to puzzle the common folk than to get their universal assent (not that they are the final arbiters of truth). An extreme case of this one in which *X* concerns some tiny, localized event, and *Y* concerns some enormous widespread event, both of which actually occur. According to centering, ‘If I had blinked just now, the entire history of the universe would have been ___’ comes out true if we insert into the blank a true statement of the entire history of the universe, for in fact I *did* blink just now.

Secondly, matters are worse when there *is* a connection between *X* and *Y*, but of the wrong sort: when the antecedent, if anything, tends to prevent or inhibit the consequent, but despite the truth of the antecedent, the consequent still manages to be true. Then *X* is evidence *against Y*. Consider the 1989 Australian Rules Football Grand Final between Geelong and Hawthorn. (It helps the example if you don’t know what happened; better still if you don’t even know what Australian Rules Football *is*.) I tell you: “If Geelong had completely outplayed Hawthorn in the final quarter, they would have won.” I expect you will want to read ‘they’ as referring to Geelong in order to render this counterfactual true. But according to centering, it is true iff we read ‘they’ as referring to Hawthorn. For as things actually turned out, Geelong *did* completely outplay Hawthorn in the final quarter, and despite that Hawthorn *did* win (their three-quarter time lead proved to be unassailable).

Thirdly, according to centering, X & Y entails not just $X \Box\!\!\rightarrow Y$, but also $Y \Box\!\!\rightarrow X$. The conjunction of the two counterfactuals is a biconditional—we might call it a *bicounterfactual,* and symbolize it:

$X \leftarrow\!\!\Box\!\!\rightarrow Y$.

The previous two problems are only exacerbated. For example, relations of counter-support often go in both directions. ‘If Gore had won the popular vote in the 2000 election, then Bush would have won the election’ is a case in point. Yet by centering, it is true that

Gore wins the popular vote $\leftarrow\Box\!\!\rightarrow$ Bush wins the election.

Fourthly, centering faces a problem with ‘might’ counterfactuals, assuming them to be duals of ‘would’ counterfactuals. ‘If I had tossed the coin, it might have landed tails.’ Arguably, this is true even if, in fact, I did toss the coin and it landed heads. But by centering, the ‘would’ counterfactual is true, so this ‘might’ counterfactual is false.

Finally, centering commits us to a dubious inequality concerning the probability of a counterfactual:

$$P(X \Box\!\!\rightarrow Y) \geq P(X \,\&\, Y),$$

(a special case of the theorem of probability theory that $P(U) \geq P(V)$ whenever V implies U). The inequality is sufficiently dubious, indeed, that *Lewis himself* seems to doubt it (1986, 22). He gives an example involving a chancy counterfactual of the form $\mathrm{A} \Box\!\!\rightarrow \mathrm{C}$, and he concludes that $\mathrm{P}(\mathrm{A} \Box\!\!\rightarrow \mathrm{C}) \approx 0$. He appeals to much the same intuition as the one I began with regarding coin tossing: since $\mathrm{A} \Diamond\!\!\rightarrow \neg\mathrm{C}$ is very probably true, $\mathrm{A} \Box\!\!\rightarrow \mathrm{C}$ is very probably false. Yet in Lewis’s example, $P(A \,\&\, C)$ may be as high as 0.97.

Thus, I am not convinced that I must automatically grant the truth of all counterfactuals with true antecedents and consequents. In fact, I’m tempted to say that those that are true are—you guessed it—those with some sort of necessary connection between antecedent and consequent, in which case we have just more examples of case 6.1.3. Then the true counterfactuals are a very special breed indeed—they are really those whose truth is secured by corresponding strict conditionals!

But even if I concede that centering secures the truth of some counterfactuals, I still think they represent a very small proportion of the counterfactuals that we actually utter. Much of the *point* of uttering counterfactuals, after all, is to convey information about a hypothetical scenario, one known or believed or assumed to be non-actual. If for nothing but Gricean reasons, we usually don't assert '*X* □→ *Y*' if we know or believe or assume *X* and *Y* both to be true; we simply assert their conjunction instead, which (assuming centering) is more informative. To be sure, we sometimes accommodate an interlocutor who disagrees with us on the truth of *X* by asserting the less informative counterfactual. But this often happens only as a concession *after* disagreement over the conjunction has emerged; and often it doesn't even happen then.

*　　*　　*　　*　　*　　*　　*

That's where my concessions end—and some of them were not really concessions after all. Even granting as true all of the counterfactuals in §§6.1 – 6.5, it is somewhat disquieting that they were all underwritten by corresponding strict conditionals or conjunctions—none of them seem to get to the heart of *counterfactuality*. All other counterfactuals, I claim, are false, and they form the vast majority. And so I conclude again, now I hope with even more justification than before: *most counterfactuals are false*.

7. Rethinking the logic of counterfactuals

Certain argument forms, which are valid for the material and the strict conditional, are said to be invalid for the counterfactual. Here are some examples:

Transitivity:

A □→ B

$\underline{B \mathrel{\Box\!\!\rightarrow} C}$

$\therefore A \mathrel{\Box\!\!\rightarrow} C$

Stalnaker (1968) gives the famous counterexample:

If J. Edgar Hoover had been born a Russian, then he would have been a Communist

If he had been a Communist, he would have been a traitor.

∴ If he had been born a Russian, he would have been a traitor.

Here, supposedly the premises are true but the conclusion false. Or consider

Strengthening the antecedent:

$\underline{A \mathrel{\Box\!\!\rightarrow} C}$

$\therefore (A\&B) \mathrel{\Box\!\!\rightarrow} C$

Lewis (1973) gives the following counterexample:

If I had struck that match, it would have lit.

∴ If I had struck that match, and it had been soaking in water, it would have lit.

Again, supposedly the premise is true but the conclusion is false.

Contraposition also putatively fails for counterfactuals; and so on. See Lewis (1973, §1.8) for more examples and discussion.

Contra Stalnaker, Lewis, and apparently received wisdom, I claim that we don't really have counterexamples to these inference patterns, because the premises are in fact not true. So there is no evidence from the failure of these patterns that counterfactuals obey some special logic.

Or consider again another putative logical principle governing counterfactuals, *conditional excluded middle:*

$$\text{(CEM)} \qquad (A \mathrel{\Box\!\!\rightarrow} C) \vee (A \mathrel{\Box\!\!\rightarrow} \neg C)$$

Stalnaker subscribes to (CEM) – indeed, its status is the chief point of disagreement between him and Lewis. But by my lights, (CEM) is no logical principle, and indeed it fails for all cases except those conceded in the previous section. For example, 'if I were to jump, I would come down' is false (because I might not come down), and 'if I were to jump, I would not come down' is false (because I might not come down), and 'if I were to jump, I would not come down' (because I might come down). Thus both disjuncts, and hence the disjunction, is false:

(I jump □→ I come down) v (I jump □→ ¬ I come down)

This should come as no surprise, given my argument that all true counterfactuals are underwritten by either strict conditionals or conjunctions. For (CEM) is false if we replace the '□→'s by either fish-hooks, ampersands, or one of each.

If I'm right about this, it undercuts an argument in favor of similarity-based accounts: the fact that they give an elegant account of the failures of these inference rules.

8. How is our practice of uttering counterfactuals vindicated?

And yet we go on cavalierly uttering counterfactuals. How, then, does our practice survive?

Here is my best hypothesis: In the neighbourhood of the ordinary but false counterfactuals that we utter, there are closely related counterfactuals that are true but not ordinary. They are counterfactuals with appropriate probabilistic or imprecise consequents. Where the consequents are probabilistic, "appropriate" means: the chance stated in the consequent is sufficiently high. If the chance is sharp, the consequent states it correctly or implies it. If the chance is imprecise, the consequent states the region of imprecision correctly or implies it.

Where the consequents are not probabilistic, “appropriate” means: the antecedent is sufficiently precise to accommodate the precision of the consequent. In each case, the counterfactual is made true by a corresponding strict conditional.

We are guilty, then, of what we might call *rounding errors*—we treat high chance propositions as certainties, low chance propositions as impossibilities, within the scope of these counterfactuals. But this is a minor crime, at least when committed by the folk in ordinary contexts. (It may not be so minor when committed by philosophers in the extraordinary contexts that they create—more on that shortly.) It is reasonable for us to do so when the rounding errors are small, as they often are. We say, for instance: ‘if I were to jump, I would come down’, which strictly speaking is false; but doing so may be justified pragmatically by the truth of a neighboring counterfactual, as it might be: ‘if I were to jump, I would come down with chance 0.99999993’.

Or perhaps, as I argued earlier, this isn’t quite right either, for I gave some reasons for skepticism even about counterfactuals with probabilistic consequents. So perhaps our practice is vindicated by counterfactuals that combine two of the devices that I offered as helping to secure their truth, probability, and imprecision:

‘if I were to jump, I would come down with chance 0.99999993 ± 0.00000001’;

or more imprecise still:

‘if I were to jump, I would come down with chance at least 0.97;

or even more imprecise still:

‘if I were to jump, I would come down with very high chance’;

So there are true counterfactuals closely related to the ones we assert that support our practice, at least when the prevailing standards for asserting counterfactuals are somewhat forgiving, as they typically are on the street. So we can legitimately assert various counterfactuals. Still, most of them remain false.

9. Is this merely philosophical pedantry?

You may grant all this, but not be disturbed by it. You may say that we already knew that various things we say are not strictly speaking true, but they are close enough to true to serve our purposes well enough. As Unger (1975) would argue, according to strict philosophical standards, nothing would count as *flat*. Kansas is not *really* flat—why, it varies in elevation by several feet! The most carefully crafted pool table is not really flat—why, even the naked eye can discern tiny bumps in the felt! One could point out the lack of flatness of Kansas, or the table, but in most contexts doing so would simply be tiresomely pedantic. Speaking the way we do about their flatness serves our purposes well enough. So you may say that I'm just being tiresomely pedantic here. Our rounding errors are as harmless as saying "It's 2 o'clock" when your watch reads 1:58.

So counterfactuals with probabilistic consequents support our practice, at least when the prevailing standards for asserting counterfactuals are somewhat forgiving, as they typically are on the street. Or so you may say.

My reply has three parts: the street is not always forgiving; even when it is, falsehood is merely forgiven rather than eradicated; and we are not always on the street. Earlier I attempted some damage control. Now it is time for some:

10. More damage

10.1 The street is not always forgiving

Even the person on the street is not immune to the effects that I have highlighted, for even on the street contexts can be created in which improbable possibilities are salient. I have not bought a ticket in this week's million-ticket lottery. I say: "If I were to buy a ticket, it would lose". You can make me take that back by having me concede that there's a one-in-a-million chance that any given ticket wins, and thus if I were to buy a ticket, it *might* win. After all, why else do people buy tickets? The very act of buying a ticket makes salient a possibility that is very improbable: *this very ticket's winning*.

And even on the street, we can make bizarre possibilities salient. Hollywood could easily make a movie whose plot turns on various characters quantum tunneling—in fact, I'm surprised that as far as I know, such a movie hasn't been made already! Having been told about quantum tunneling, you cannot immediately ignore it. A context has been set up in which quantum tunneling is a live, salient possibility. In that context, it is not tiresome pedantry to balk at various ordinary counterfactuals whose truth is sabotaged by the possibility of quantum tunneling. On the contrary, it is required of a cooperative, attentive audience.

10.2 On the street, falsehood is merely forgiven, not eradicated

Granted, on the street we get away with uttering various counterfactuals whose truth I have questioned. Still, that is no proof that they *become true* on the street. Earlier I expressed my doubts about a contextualist account according to which they do. And the linguistic data that we have about street-talk could equally be explained by an account according to which the counterfactuals become *assertable*. I suggest, then, that high probabilities of their consequents that fall short of 1 suffice for assertability, but not for truth. How high the probabilities need to be may be context-dependent, in a way that the falsehood of the counterfactuals is not. I can thus concede that much to the friends of context-dependence regarding counterfactuals, without conceding the context-dependence of their truth-values.

I think there is an important disanalogy here between 'flatness' and counterfactuals—and here I part company with Unger. I think 'flat' (and indeed all other gradable adjectives) is tacitly *comparative*. While it looks like a one-place property, it is really a two-place relation. When we say '*x* is flat', we mean '*x* is flatter than __', where the '__' is filled in by a salient alternative to *x*—perhaps a neighbouring case to *x*, or the average of a set of cases to which *x* belongs. Thus, when I say that Kansas is flat, I may be comparing its degree of overall flatness to the surrounding landscape, or to other states. Moreover, once context makes clear the second relatum, what I say may be *true*—pace Unger. Flatness of landscapes comes in degrees; relative to some alternatives a given landscape may have a higher degree of flatness, while relative to others it has a lower degree.

But I don't think that truth comes in various degrees—pace degree-of-truth theorists. I can make no sense of the notion that a given counterfactual is 'truer than' another, while 'less true than' a third one. Truth comes in two degrees. A miss is as good as a mile; and most counterfactuals miss.

So at best, various counterfactuals that we utter on the street—which are almost always ordinary—are assertable (although of course many of them are not even that). And as followers of Adams have emphasized regarding indicative conditionals, assertability does not imply truth. He went further, and thought that indicative conditionals don't have truth values at all—not that that stops us from uttering them. I, by contrast, am happy to grant that counterfactuals *have* truth values. In many cases, however, they are not what we offhand took them to be—not that that stops us from uttering them.

10.3 We are not always on the street

I have claimed that we are guilty of various tiny rounding errors, ignoring the tiny probabilities that under various counterfactual suppositions, various consequents turn out false. (Sometimes this is the even tinier rounding error of treating a possibility with zero probability as if it is impossible.) Moreover, I agree with the spirit of 'your' objection in §9: perhaps it *is* tiresomely pedantic to worry about tiny rounding errors. But it is a philosopher's job to be tiresomely pedantic. And I don't just mean that while I'm presenting a philosophical work such as this, I can reasonably raise standards of precision above what they would be on the street. That's true—unlike most counterfactuals—but not all that surprising, especially in these post-Ungerian, post-Lewisian times. The problem runs deeper than that. *For various philosophers employ counterfactuals in their philosophical positions, or in their conceptual analyses*. Counterfactuals are, as I said at the outset, a philosophical staple these days: they figure in influential analyses of causation, perception, knowledge, personal identity, laws of nature, rational decision, confirmation, dispositions, free action, explanation, and so on. There, the standards of precision are high, the context unforgiving. And when they are high, philosophers can't so easily plead the person on the street's excuse. The danger is that the various counterfactuals that are supposed to underpin causation, perception, knowledge, personal identity, … turn out to be false, instead of true, when their analyses *require them to be true*. (We can safely assume that these counterfactuals are not on my list of concessions in §6, although you may like to check that for your favourite counterfactual-laden analysis.) By analogy, if the analysis of some concept dear to us required some things to be *perfectly flat,* and it turned out that there were no such things, then either the analysis, or the concept, would be in trouble. (This last sentence, while a counterfactual, is trivially true.) Fortunately, we do not run into this problem with flatness.

However, we *often* run into this problem with counterfactuals. Either the various philosophical analyses that appeal to them are mistaken, or we live in a world devoid of causation, devoid of perception, devoid of knowledge, devoid of persisting persons, and so on. (This is an inclusive 'or'.) Assuming that our world is not so impoverished, it is the analyses that are under threat.

In §5.1, I countenanced the views that the counterfactuals I was considering were *indeterminate* or *lack truth values* rather than being false. I argued against those views, but let me grant each of them in turn, now, for the sake of the argument. Rewrite this section, if you like, replacing the word "false" by "indeterminate" or "lacking in truth value". This should hardly comfort the relevant philosophers, and it will hardly save their analyses. For the philosophers think that various claims of causation, of perception, and so on, are *true*. This is not the case if their analyses are right, and if the counterfactuals that figure in those analyses are indeterminate or lack truth value. So it is not true that we live in a world with causation, perception, knowledge, persisting persons, and so on. Either our commonsense beliefs about the world require wholesale revision, or, more likely, these philosophical positions do.

Nor did I feel I had to allow that the truth values of most counterfactuals are context-dependent (§5.2), but let us grant that now for the sake of the argument. Then the counterfactuals that figure in the analyses of causation, perception, knowledge, personal identity, and so on are context-dependent. This in turn means that if these analyses are correct, or at least correct insofar as they employ counterfactuals (even if their exact details are incorrect), the analysanda are context-dependent, unless somehow the context-dependences 'cancel out'.[20] To see how this is possible, suppose we analyze 'even number' as 'either a large number divisible by 2 or a non-large number divisible by 2. Each disjunct of the analysis is context-dependent, since 'large' is. However, the context-dependence of each disjunct compensates for the other, taking up its slack: as we move to a context that is more demanding for 'large', we *ipso facto* move to a context that is less demanding for 'non-large', and vice versa. As a result, 'even number' proves not to be context-dependent after all. Good news for 'even number'! However, I very much doubt that there will be such good news for counterfactual analyses of 'causation', 'perception' and so on. While we would need to run through these analyses on a case-by-case basis, I bet that in each case the context-dependence of their counterfactuals won't cancel out.

The threat, then, is that if these analyses are correct, then much that we hold dear turns out to be context-dependent. Perhaps in some cases that may not be such a worry; in fact, perhaps we already had reason to believe it. Contextualism about some those concepts already has some currency—see e.g. Schaffer (2006) on causation, or Lewis (1996) on knowledge, or van Fraassen (1980) on explanation. Note, however, that the context-dependence of counterfactuals may reveal a *further*, perhaps unintended, unwanted, and hitherto unappreciated source of context-dependence in these concepts. Moreover, contextualism is rather less appealing for some of the other concepts—for example, for laws of nature, or perception, or personal identity. Imagine a defence lawyer telling the jury that it is *context-dependent* whether the defendant is the same person as the murderer! And do you think that

[20] I thank Alex Byrne for this observation.

whether or not you are the same person as you were when you began reading this manuscript is, despite its considerable length, *context-dependent*?!

To be sure, these philosophical analyses are given *in a particular context*—a philosophical context. That may alleviate the problem of context-dependence: once and for all, the counterfactuals in the analyses should be evaluated by the standards of that context. But this only makes worse the problem of the *falsehood* of the relevant counterfactuals. After all, philosophical contexts are demanding: bizarre possibilities are fair game. We can legitimately entertain possibilities in which billiard balls quantum tunnel to China or in which human bodies vaporize; indeed, if we are doing our job properly, we should be so imaginative. This brings us back full circle to my argument for the falsehood of most 'would' counterfactuals from the truth of corresponding 'might not' counterfactuals—such were the possibilities that I entertained when arguing for the truth of the various 'might not' counterfactuals that undermined the corresponding 'would' counterfactuals. I gave that argument in a particular context—a philosophical context.

Well may we wonder, then, how the slogan "counterfactuals are context-dependent" can be repeated so blithely when philosophers so often reach for counterfactuals in their conceptual analyses. Or perhaps we should wonder, rather, how philosophers can reach so blithely for counterfactuals when that slogan is repeated so often! Either way, something is amiss. I can't accept that the facts about the laws of nature, or perception, or personal identity, are context-dependent. Nor can I accept that we should be eliminativists about these things because our favorite philosophical analyses of them involve counterfactuals that are false. I suspect, rather, that the fault lies with the analyses. Again, this should really be addressed on a case-by-case basis, and our verdicts may differ across the cases. Still, I hope that *my* slogan regarding counterfactuals—that most of them are false—will alert us to the dangers that one courts when trafficking in them while philosophizing. A similar caution carries over to their use in the sciences and the social sciences.

* * * * * * *

If had more time, I would say still more about all of this. But that's a counterfactual, and it's false.[21]

School of Philosophy
Research School of Social Sciences
Australian National University
Canberra, ACT 0200
Australia

[21] My greatest debt is to David Lewis–anyone familiar with his work should be able to see how much it has influenced me. I wrote a short first draft of this paper when I was a visiting fellow in 2002 at the Center for Philosophy of Science at the University of Pittsburgh, which provided a very stimulating research environment. My thanks to all who helped me there. I am grateful to many people over the years since then for helpful discussions and/or comments on this material that led to improvements therein, especially: Robert Adams, Karen Bennett, Eliza Block, Rachael Briggs, Alex Byrne, Jim Cargile, David Chalmers, Fiona Cowie, Vladan Djordjevic, Jamie Dreier, David Etlin, Daniel Greco, Kenny Easwaran, Antony Eagle, Dorothy Edgington, Andy Egan, Adam Elga, Justin Fisher, Hilary Greaves, Allen Hazlett, Chris Hill, Chris Hitchcock, Carrie Jenkins, Jenann Ismael, Martin Jones, Jeff King, Boris Kment, Dan Korman, Douglas Kutach, Igal Kvart, Marc Lange, Stephan Leuenberger, Barry Loewer, Aidan Lyon, John McFarlane, Sarah Moss, Daniel Nolan, Denis Robinson, Joe Salerno, Paolo Santorio, Jonathan Schaffer, Wolfgang Schwarz, Ted Sider, Nico Silins, Declan Smithies, Robert Stalnaker, Jason Stanley, Brian Weatherson, Duncan Watson, Lee Walters, John Williams, Jim Woodward, Lyle Zynda, and audiences at Indiana University South Bend, Melbourne University, Monash University, Singapore Management University, University of New South Wales, Oberlin College, Arizona State University, University of North Carolina at Chapel Hill, McGill University, the December 2006 meeting of the APA in Washington, the 2007 meeting of the AAP in Armidale, Oxford University, Leeds University, Washington University, Brown University, Yale University, Dartmouth University, Princeton University, Rutgers University, University of Connecticut at Storrs, University of Arizona, Oslo University, Katholieke Universiteit Leuven and the participants at the Summer School on Conditionals in Budapest in 2009. [**This list is old and needs a considerable amount of updating.**] If this manuscript had not benefited from their comments, it would be poorer.

REFERENCES (to be continued and filled in)

Adams, Ernest (1965): "The Logic of Conditionals", *Inquiry* 8, 166-197.

Adams, Ernest (1975): *The Logic of Conditionals*, Dordrecht: Reidel.

Adams, Robert Merrihew (1977): “Middle Knowledge and the Problem of Evil”, *American Philosophical Quarterly* 14, No. 2, 109 – 117.

Arntzenius (200x):

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

Bigelow, John and Robert Pargetter (1990): *Science and Necessity*, Cambridge: Cambridge University Press.

Borel, Emil

DeRose, Keith (1999): “Can It Be That It Would Have Been Even Though It Might Not Have Been?”, *Philosophical Perspectives* 13, Epistemology, 385 -413.

Eagle (200x)

Edgington, Dorothy

Elga, A. (2004). ‘Infinitesimal chances and laws of nature’. Australasian Journal of

Philosophy, 82, 67–76.

Elgin, Catherine (1988): “The Epistemic Efficacy of Stupidity.” *Synthese* 74: 297-311.

Fine, Kit (1975): Critical notice: Counterfactuals. Mind 84: 451–58.

Fitelson, Branden (20xx): “Strengthening the Case for Knowledge from Falsehood”, *Analysis*.

Hájek, Alan (1997): "'*Mises Redux*'—*Redux:* Fifteen Arguments Against Finite Frequentism", *Erkenntnis*, Vol. 45, 209-227. Reprinted in *Probability, Dynamics and Causality – Essays in Honor of Richard C. Jeffrey,* D. Costantini and M. Galavotti (eds.), Kluwer.

Hájek, Alan (2003): "What Conditional Probability Could Not Be", *Synthese,* Vol. 137, No. 3, 273-323.

Harper, W. L., R. Stalnaker, and G. Pearce (eds.) (1981): *Ifs*, Dordrecht: Reidel.

Halliday and Resnick (19xx): Fundamentals of Physics, Vol. 2, John Wiley & Sons.

Hawthorne, J. (2005). 'Chance and counterfactuals'. Philosophical and Phenomenological Research, pages 396–405.

Heller, Mark (19xx)

Hoefer (forthcoming)

Jackson (19xx)

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

Jeffrey, Richard (1977): "Mises Redux", reprinted in (1992): *Probability and the Art of Judgment*, Cambridge Studies in Probability, Induction, and Decision Theory.

Kment, Boris

Kolmogorov, A. N.

Levi (19xx)

Loewer (2001).

Lewis, David (1969): *Convention: A Philosophical Study,* Harvard University Press.

Lewis, David (1973): "Counterfactuals and Comparative Possibility", *Journal of Philosophical Logic* 2; reprinted in Lewis 1986.

Lewis, David (1973b): *Counterfactuals*, Basil Blackwell.

Lewis, D. K. (1979). ‘Counterfactual dependence and time’s arrow’. Noˆus, 13, 455–76. Reprinted with postscript in Lewis, Philosophical Papers II (Oxford University Press, 1986) 32–51. Also reprinted in Jackson (ed) Conditionals (Oxford University Press, 1991) 46-76.

Lewis, David (1986): *Philosophical Papers,* Vol. II, Oxford University Press.

Lewis, David (1994): “Humean Supervenience Debugged”, *Mind* 103, 473-490.

Lewis, David (1996): “Elusive Knowledge”, *Australasian Journal of Philosophy* 74, No. 4, 549-567.

Mates, B. *Stoic Logic* (Berkeley & Los Angeles: University of California Press, 1961), 43.

Nute, Donald

Plantinga, Alvin (1974): *The Nature of Necessity,* Oxford University Press.

Pollock, John (1976): *Subjunctive Reasoning*, Boston: Reidel.

Quine

Schaffer, Jonathan (2006): “Contrastive Causation”, *Philosophical Review*.

Schaffer, Jonathan: Counterfactuals, causal independence and conceptual circularity

Shafer, Glenn (2006): “From Cournot’s Principle to Market Efficiency”.

Sorensen, Roy (1988): *Blindspots,* Oxford University Press.

Stalnaker, R. (1968): "A Theory of Conditionals", in N. Rescher (ed.), *Studies in Logical Theory*, Blackwell.

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

Unger, Peter (1975): *Ignorance: A Case For Skepticism,* Oxford: Clarendon Press.

van Fraassen, Bas (1980): *The Scientific Image,* Oxford: Clarendon Press.

van Inwagen (19xx).

Warfield, Ted (2005): “Knowledge from Falsehood”, *Philosophical Perspectives* 19, 405–16.

Press, Cambridge.

Williams, R "Chances, counterfactuals and similarity"; *Philosophy and Phenomenological Research*, vol 77(2), 2008.

Williamson, Timothy