Synthese (2013) 190:639–660
DOI 10.1007/s11229-012-0196-6

Remarks on counterpossibles

Berit Brogaard · Joe Salerno

Received: 15 August 2012 / Accepted: 15 August 2012 / Published online: 21 December 2012

Abstract Since the publication of David Lewis' Counterfactuals, the standard line on subjunctive conditionals with impossible antecedents (or counterpossibles) has been that they are vacuously true. That is, a conditional of the form 'If p were the case, q would be the case' is trivially true whenever the antecedent, p, is impossible. The primary justification is that Lewis' semantics best approximates the English subjunctive conditional, and that a vacuous treatment of counterpossibles is a consequence of that very elegant theory. Another justification derives from the classical lore than if an impossibility were true, then anything goes. In this paper we defend non-vacuism, the view that counterpossibles are sometimes non-vacuously true and sometimes non-vacuously false. We do so while retaining a Lewisian semantics, which is to say, the approach we favor does not require us to abandon classical logic or a similarity semantics. It does however require us to countenance impossible worlds. An impossible worlds treatment of counterpossibles is suggested (but not defended) by Lewis (Counterfactuals. Blackwell, Oxford, 1973), and developed by Nolan (Notre Dame J Formal Logic 38:325–527, 1997), Kment (Mind 115:261–310, 2006a: Philos Perspect 20:237–302, 2006b), and Vander Laan (In: Jackson F, Priest G (eds) Lewisian

To appear in J. van Bentham, V. Hendricks, J. Symons, and S. A. Pedersen (eds.) *Between Logic and Intuition: David Lewis and the Future of Formal Methods in Philosophy*, Synthese Library, 2008.

B. Brogaard (✉) · J. Salerno
Australian National University, Canberra ACT, Australia
e-mail: brogaardb@gmail.com

B. Brogaard
University of Missouri, St. Louis, MO, USA

J. Salerno
St. Louis University, St. Louis, MO, USA

themes. Oxford University Press, Oxford, 2004). We follow this tradition, and develop an account of comparative similarity for impossible worlds..

Keywords Impossible worlds · Accessibility relation · Counterfactuals · Subjunctive conditionals · Indicative conditionals · Lewis semantics · Non-trivial counterpossibles · Limit assumption · Essential properties · Contingent counterpossibles

Since the publication of David Lewis' *Counterfactuals,* the standard line on subjunctive conditionals with impossible antecedents (or counterpossibles) has been that they are vacuously true. That is, a conditional of the form 'If p were the case, q would be the case' is trivially true whenever the antecedent, p, is impossible. The primary justification is that Lewis' semantics best approximates the English subjunctive conditional, and that a vacuous treatment of counterpossibles is a consequence of that very elegant theory. Another justification derives from the classical lore than if an impossibility were true, then anything goes.

In this paper we defend *non-vacuism*—the view that counterpossibles are sometimes non-vacuously true and sometimes non-vacuously false. We do so while retaining a Lewisian semantics—which is to say, the approach we favor does not require us to abandon classical logic or a similarity semantics. It does however require us to countenance impossible worlds. An impossible worlds treatment of counterpossibles is suggested (but not defended) by Lewis (1973), and developed by Nolan (1997), Kment (2006a,b), and Vander Laan (2004). We follow this tradition, and develop an account of comparative similarity for impossible worlds.

1 A Lewisian theory of counterfactuals

Lewis' account is usually abbreviated as follows.

> (Account 1) A subjunctive, 'if it were the case that p, it would be the case that q', is true just when every closest p-world is a q-world, where 'closest' means 'most similar (to the world of utterance)'.

It immediately follows that a subjunctive with an impossible antecedent is vacuously true. For if there are no p-worlds, then, vacuously, every closest p-world is a q-world.

More carefully, Lewis' account tells us that

> (Account 2) A subjunctive conditional, 'p implies q', is true just in case either (i) there is no p-world, or (ii) there is a $\{p, q\}$-world that is closer than any $\{p, \sim q\}$-world.

By definition counterpossibles satisfy clause (i), and so, are vacuously true.

The difference between (Account 1) and (Account 2) becomes evident when we deny the limit assumption. The *limit assumption* tells us that there is a limit on how close the possible worlds can be. It says whenever there is a p-world, there is at least one closest p-world. Lewis rightly denies the assumption. $w1$ is closer than $w2$ to the actual world just in case $w1$ is "more similar" to the actual world than is $w2$. In this context to deny the limit assumption is to deny that there is always a limit on how similar to the actual world the possible worlds may be. However there is a possible world exactly like the actual world but where I drank a whole beer more than I in fact drank. And there is an even more similar world where I drank only a half beer more than I in fact drank, and an even more similar world where I drank only a quarter more, and so on, ad infinitum. These worlds become more and more similar to the actual with respect to how much alcohol was consumed, because the amount consumed gets closer and closer to the actual amount. Now suppose my blood alcohol content (b.a.c.) is .08 and that legally one cannot drive if one's b.a.c. is any higher. Then, by (Account 2), (a) comes out true and (b) comes out false:

(a) If I had had any more to drink, my b.a.c. would have been over the limit.
(b) If I had had any more to drink, my b.a.c. would NOT have been over the limit.

(a) is true on (Account 2) because there is at least one {more-to-drink, b.a.c.-over- the-limit}-world that is closer than any {more-to-drink, b.a.c.-NOT-over-the-limit}-world. (b) is false on that same account because there is no {more-to-drink, NOT-over-the-limit}-world that is closer than every {more-to-dink, over-the-limit}-world. On (Account 1), by contrast, both (a) and (b) come out vacuously true—but not because there is no possible world where I had more to drink, but rather because there is no *closest* possible world where I had more to drink. There are always closer and closer worlds where I had more to drink.

The account we prefer says this:

(Account 3) A subjunctive, 'if it were the case that p, it would be the case that q', is true just when every closest p-world is a q-world, where 'closest' is read as 'most relevantly similar'.

The account has the advantage of making explicit something that Lewis and others assume when utilizing a Lewisian theory of counterfactuals, viz., 'comparative similarity is vague' and 'context is needed to resolve some of the vagueness and to determine which similarities and differences are important' (Lewis 1973, p. 67 and 1986, p. 42).

Notice that, unlike (Account 1), our latest formulation does not treat sentences (a) and (b) as both vacuously true. The {more-to-drink}-worlds that are most relevantly similar to the actual world are all of the possible worlds where I had something more to drink. A world where I have another pint is as relevant for the evaluation of the conditional as a world where I have another half, quarter, eighth, or sixteenth of a pint. In all those worlds, my b.a.c. goes over the limit. So (a) is true. In none of those worlds does my b.a.c. fail to go over the limit. Hence (b) is false.

(Account 3) helps to avoid other counterexamples to Lewis' theory (Lewis 1973, pp. 20–21 and Lewis 1981, pp. 228–230). Consider the claim

If I were at least 6 feet tall, then I would be exactly 6 feet tall. (Hájek, manuscript)

The claim is false, since if I were at least 6 feet tall I might be 6 foot one. But, since in fact I am less than 6 feet tall, a world where I am exactly 6 feet is more similar than any world where I am more than 6 feet tall. Hence, on Lewis's favored reading (Account 2), the above claim is true. By contrast, on (Account 3), the sentence is treated properly as false. And that's because, in an ordinary context where the above conditional is asserted, there is a range of possible heights that matter and worlds where I am six-foot-one are as relevant as worlds where I am precisely six-foot. With respect to height, worlds where I am precisely six-foot may be more similar than worlds where I am taller than six-foot, but they needn't be more relevantly similar.

Considerations of relevance are precisely the kinds of considerations that Lewis takes into account when dealing with the oddity of the Caesar sentences:

(1) If Caesar commanded in Iraq, he would have used nukes.
(2) If Caesar commanded in Iraq, he would have used catapults.

In a conversational context that makes salient Caesar's ruthlessness, we find that (1) is true and (2) is false. In an altogether different context that makes salient the state of military technology at the time of Caesar's reign, we find that (2) is true and (1) is false. The key to a solution is to focus on the vagueness of comparative similarity. Context is needed to determine which similarities and differences are important. Our explicit mention of relevance is meant to preserve this Lewisian insight. On either reading of the Caesar sentences, it is the most relevantly similar {Caesar-commands-in-Iraq}-worlds that matter for the proper evaluation of the counterfactual.

A notion of closeness in terms of relevant similarity allows one to bypass a family of objections to Lewis' account that turn on the vagueness of comparative similarity. Moreover, and perhaps for that very reason, something like (Account 3) has entrenched itself in the mainstream philosophical understanding of counterfactuals.

(Account 3) provides us with our starting point. We'll have it in the background when we talk about the "standard account". Like Accounts 1 and 2, (Account 3) treats all counterpossibles as vacuously true. If there are no antecedent worlds, then there are no most relevantly similar antecedent worlds, in which case, it is vacuously true that all the most relevantly similar antecedent worlds are consequent worlds.

2 Some virtues of non-vacuism

In this section we argue that vacuism (i) fails to give the correct truth-conditions for some counterpossibles, and (ii) fails to explain the invalidity of some invalid arguments that embed counterpossibles. Further, non-vacuism is better equipped (i) to explain the informativeness of counterpossibles philosophy, and (ii) to draw a modal distinction between essence and de re necessity.

2.1 Vacuism fails to give the correct truth-conditions

Counterpossibles are trivial on the standard account. By 'trivial', we mean *vacuously true and semantically uninformative*. Counterpossibles are *vacuously true* in that they are always true; an impossibility counterfactually implies anything you like. And relatedly, they are *uninformative* in the sense that the consequent of a counterpossible makes no contribution to the truth-value, meaning or our understanding of the whole. But the claim that counterpossibles are trivial runs counter to our intuitions. Consider:

(1) If Hobbes had (secretly) squared the circle, sick children in the mountains of South America at the time would not have cared. (Nolan 1997, p. 544)

(2) If intuitionistic logic were the correct logic, then the law of excluded middle would still be unrestrictedly valid.
(3) If it had actually been raining (here and now), then it would not have been raining (here and now).

The intuition is that (1) is true, but non-vacuously. Sick children in South American could not care less about the intellectual achievements of faraway philosophers, especially when said achievements are produced in secret. (2), on the other hand, is false, since by definition intuitionistic logic rejects the unrestricted validity of excluded middle.[1] (3) is a more subtle kind of counterpossible. Since in fact it is not actually raining (here and now), it follows that in every possible world it is false that in the actual world it is raining (here and now). However, (3) is non-vacuously false. If it had actually been raining then indeed it would have been raining! In each of these examples, the antecedent is impossible and the consequent seems critical to the truth-value of the whole. But, if vacuism is true, the consequent of each is irrelevant, and all three examples are vacuous and uninformative. That is one piece of evidence against the standard account and vacuism more generally.

2.2 Vacuism fails to explain the invalidity of some clearly invalid arguments

Consider the following scenario. An overly enthusiastic grad student believes his counterexample would have refuted Fermat's last theorem (FLT), if only the theorem were false. His advisor disagrees since there is no telling which counterexample would have refuted FLT, if it were false. The student complains that the advisor is discriminating against him on the grounds that he is a lowly graduate student. The advisor reassures the student that his not having yet completed grad school has nothing to do with it. He says, 'Even if you complete grad school, it still will be false that you would have refuted FLT if FLT were false'. By the standard logic, the director proceeds from there to prove all too easily that the student will never complete grad school.[2]

(Enthusiastic Grad Student)
(i) (Even) if you finish grad school, it (still) will be false that *if Fermat's last theorem had been false, you would have been able to prove it false.*
(ii) Necessarily, Fermat's last theorem is true.

So,

(iii) You won't finish grad school.

The teacher reasoned as follows. Necessarily FLT is true. So it's impossible that it's false. By vacuism, had it been false you would have been able to prove it false. This

[1] We assume (2) is a counterpossible. Ex hypothesi, classical logic is the correct logic, and so, by reasonable assumptions, is correct necessarily. Hence, it is impossible that intuitionism is correct. If indeed intuitionistic logic is the correct logic, the example can obviously be modified to suit this fact.

[2] Counterfactuals are often said to defy many classically valid inference schemas. However, the schema appealed to here ought to be valid. Moreover, a strong argument can be made against the claim that counterfactuals do violate many classically valid inference schemas. See Brogaard and Salerno (2008).

contradicts the consequent of premise i. So he denies the antecedent, and concludes that the student will never finish grad school. The derivation appears formally below:

1.	$\bullet$FLT	Premise ii
2.	$\neg\Diamond\neg$ FLT	from 1, by definition of $\bullet$
3.	$\neg$ FLT $\bullet \rightarrow$ P	from 2, by vacuism
4.	G $\rightarrow \neg(\neg$ FLT $\bullet \rightarrow$ P)	Premise i
5.	$\neg$ G	from 1, 4

The premises are true in the scenario described. The conclusion, however, follows all too easily. The only controversial inference is the one justified by the vacuist truth conditions for counterpossibles. Hence, vacuism cannot explain the invalidity of the above invalid argument. Let's consider further evidence against vacuism.

2.3 Vacuism fails to explain the informativeness of counterpossible philosophy

Counterpossible Philosophy isn't vacuous and uninformative, is it? Don't we meaningfully think about how things would have been with different laws of nature/metaphysics/logic? Don't we non-trivially evaluate the consequences of theories known to be impossible? For instance, we believe that if presentism had been true, there would have been no cross-time relations, and that if Plato's theory of recollection were correct, then we would have existed before we were born.[3] If we add a negation to the consequents of these claims, then we are left with something we do not believe. Indeed we believe the negations of these corresponding conditionals. In philosophy we often do treat counterpossibles '$p \Box\!\!\rightarrow q$' and '$p \Box\!\!\rightarrow \sim q$' as differing in their truth-values and content. They make distinct claims, and we understand them differently. Vacuism cannot explain this.

In sum, if all counterpossibles were trivially true, much of our substantial counterfactual philosophizing would be vacuous, uninformative and fallacious. Perhaps much of philosophy is vacuous, uninformative and fallacious. But if it is, it is not for systematic misuse of the counterfactual.

It may be urged that counterpossible talk is elliptical for something more elaborate.[4] For example, it may be suggested that when we say 'if $2 + 2$ were 5, my school teachers would have been wrong', we in fact mean 'if the meaning "2", "5" and "+" had been such that "$2 + 2 = 5$" were true, then what my school teachers actually said would have been wrong'. And when we say 'if water were not XYZ, then it would not have been H_2O', then we in fact mean 'if the watery stuff in our environment had been XYZ, then it would not have been H_2O'. This method would appear to yield the right result in a large range of cases.

The strategy, however, is unconvincing. For it is quite implausible to think that the logical form of counterpossibles is so strongly devious that in reality there are no counterpossibles. Of course, an opponent of impossible worlds might insist that we treat counterfactuals as operators on something like Fregean senses (or intensions

[3] Versions of this point are made by Nolan (1997, pp. 539–540), and Kim and Maslen (2006).

[4] Versions of the objection appear in Lewis (1973), and Wierenga (1998)

corresponding to such senses) rather than as operators on Russellian propositions (or intensions corresponding to such propositions).[5] If counterfactuals are operators on Fregean senses, then we can account for the truth of 'if water were XYZ, then it would not be H_2O' without treating the latter as a counterpossible.[6] For, the closest world where the sense of 'water is XYZ' is true is a possible world. At that world, the sense of 'water is not H_2O' is true. So, the counterfactual comes out true but not vacuously true. But this strategy, however brilliant, does not generalize. For example, there is no way to extend the Fregean approach to account for the intuitive falsehood of 'if my shirt had been red and non-red all over, then it would have been green' or 'if Hobbes had squared the circle, then sick children in South America would have cared'. For, there are no possible worlds in which the Fregean senses of 'my shirt is red and non-red all over' or 'Hobbes squared the circle' are true. We suspect that the only initially plausible explanation of our intuitions about counterpossibles, which is not simply a variation on our non-vacuist explanation, will be pragmatic. The idea underlying a pragmatic account is this: counterpossibles are vacuously true, and our intuitions that they are non-vacuously true or false can be explained away on pragmatic grounds—perhaps by offering paraphrases of the above-mentioned sort.

We offer our pragmatically-inclined opponent this reply. Simply listing paraphrases along the lines suggested above is not the same as giving a pragmatic explanation of our intuitions that counterpossibles are non-vacuously true or false. We need a pragmatic story that explains why we take counterpossibles, but not ordinary counterfactuals, to say something they do not literally say. We highly doubt that there is an elegant and convincing pragmatic story to be told. But suppose we are wrong. Suppose there is such a story. We still cannot conclude that a vacuity account plus pragmatic story is to be preferred to a non-vacuity account. The mere existence of a pragmatic story of some phenomenon does not make the phenomenon pragmatic. We need a good reason to prefer the pragmatic story to the semantic one.

Here is a related concern. One might worry that the intuitions we have about counterpossibles extend to indicative conditionals. Consider, for instance, 'if paraconsistent logic is true, then everything follows from a contradiction' and 'if $2 + 2 = 5$, then what my teachers said is correct'. Such indicative conditionals seem false for the same reason that the corresponding counterpossible subjunctive conditionals are false. They certainly do not seem true. Yet, it may be said, counterpossible indicative conditionals cannot plausibly be given a non-vacuist treatment.

We beg to differ. Though this is not the place to develop a non-vacuist account of counterpossible indicatives, we believe the very same arguments that motivate a non-vacuist treatment of counterpossible subjunctives will support a non-vacuist treatment of counterpossible indicatives. Ultimately, we think that there is no important semantic difference between indicative and subjunctive conditionals in general. But this will be the topic of another paper.

[5] An account along these lines is outlined in Brogaard (2007).

[6] If so construed, counterfactuals would be similar semantically to Chalmers' indicative conditionals (see Chalmers 1998, 2002).

2.4 Non-Vacuism falicitates a modal analysis of essence

Another, quite different, motivation for non-vacuism is that it facilitates a modal analysis of essence. Fine (1994) argues that the standard modal account of essence as de re modality is 'fundamentally misguided' (p. 3). We agree with his critique and suggest an alternative counterfactual characterization of essence. The counterfactual account lends support to non-vacuism.

Explicitly modal accounts of essence are relatively recent. Moore (1919) offers an account of essence in terms of entailment: a property F is essential (or *internal*) to x iff '$x = a$' entails 'Fx' (p. 293). Since the emergence of modal logic, essential properties are typically equated with de re necessities. But, as Fine notes, the presumption that there is 'nothing special about the modal character of essentialist claims beyond their being de re' is mistaken (p. 3). While Kripke's wooden table, Tabby, is necessarily a member of the set {Tabby}, it is not essential to Tabby that it be a member of that set. Nor is it essential to Tabby that seven is prime or that it be such that it's either raining or not. The properties: being a member of the set {Tabby}, being such that seven is prime, and being such that it's either raining or not seem irrelevant to the question of what it is to be Tabby. By contrast, the wood of which Tabby is composed seems relevant to Tabby's essence Kripke (1980).

Here is an explanation of said intuitions: if there hadn't been sets (or if seven hadn't been prime,…), then Tabby might still have existed. By contrast, Tabby wouldn't exist if there were no wood. This sort of explanation requires, for its non-triviality and informativeness, that some counterpossibles be non-trivial and informative, or more specifically, that their truth-values be affected by the truth-values of their consequents. The suggestion can be put this way. Tabby exists in some of the closest impossible worlds where there are no sets. To avoid trivialization we take impossible worlds to be *non-deductively closed* sets of sentences. We leave classical logic and one's favored modal logic intact for non-counterpossible modal discourse. We will say more later about closeness of impossible worlds. Here we simply aim to show that non-trivial counterpossibles make a modal analysis of essences possible. Or at least, it makes possible an analysis that distinguishes essence from de re necessity. The suggestion roughly is this:

> x is essentially F just in case if nothing had been F, then x would not have existed.

However, our right-to-left is curious.[7] If Mafia Mike hadn't protected his nephew Joey, then Joey wouldn't exist. Joey's enemies haven't killed him only because of Mike's protection. Yet, arguably, Joey is not essentially protected by Uncle Mike. Joey can go into witness protection, or over-ride Mikes sound advice by having all enemies systematically eliminated. In a word, it is false that Mike protects Joey in every metaphysically possible world in which Joey exists. Much worse than this apparent *problem of contingent essence* is the problem of actually *uninstantiated essence*. If there were

[7] Thanks to Mike Almeida for pressing the concern.

no medical doctors, I wouldn't exist. But I am no medical doctor. A fortiori, I'm not essentially a medical doctor.[8]

We prefer a limited tolerance for contingent essences, and recommend a technical modification to deal with the problem of uninstantiated essences. The technical modification is this. Replace 'x is essentially F' with 'F is essential to x'. Thus, there being Fs is essential to x iff if there were no Fs then x wouldn't exist. This gets around the problem of uninstantiated essential properties. For instance, if there were no doctors, indeed, I would fail to exist. By the above account, there being doctors is essential to my existence. But this does not imply that I am a doctor.

This brings us to the problem of contingent essence. The modification highlights that there are things that are contingently essential to my existence. However, the contingent nature of essence is justified by the vernacular. Lennon and McCartney once wrote, 'I won't stay in a world without love'. In some very important sense, love is essential for their continued existence. Perhaps, in the eyes of these poets, suicide is the only real option in the face of a loveless world. Or perhaps these poets couldn't go on while retaining their identity *qua* poets in a loveless world. Either way, in some ordinary sense they can't exist (or can't exist "as themselves") in a world without love. More mundane examples include:

(1) It is essential to your team's success that you advertise your website.
(2) It is essential to your productivity that you back up your hard drive.

Here 'essential' does not mean what it typically means in recent philosophical literature. (1) doesn't say that there is no world in which you don't advertise your website but your team is successful, and (2) doesn't say that there is no world in which you fail to back up your hard drive but still produce the same amount of work. It's rather something like: holding fixed relevant background conditions, if you don't advertise your website, your team will not be very successful. And holding fixed relevant conditions, if the hard drive is not backed up, work will eventually get lost and productivity will be diminished. It is natural, then, to understand ordinary essence claims as counterfactuals.

We distinguish the philosophical from the ordinary uses of 'essence' as follows.[9]

(Ordinary Use)
There being Fs is essential to x iff if there were no Fs then x wouldn't exist.
(Philosophical Use)
There being Fs is essential to x (or x is essentially F) iff (1) if there were no Fs then x wouldn't exist, and (2) it is metaphysically necessarily that if x exists then x is F

Philosophical use restricts essential properties of x to those had in every metaphysically possible world in which x exists. It has the benefit of distinguishing the essential

8 Thanks to Jim Stone for raising the objection.

9 This may not be the whole story. Ultimately, we might also want to require that F is essential to a only if: there might still be Fs even if a hadn't existed. Otherwise, we might get the result that the property of being the sum of Socrates and the number two is essential to Socrates (depending on whether one's ontology admits of the existence of such sums). However, we will leave this complication to one side in this paper.

from the necessary while ruling out the essential but contingent (and uninstantiated). On this account an essential property is not only a property that x necessarily has, but is a property that x wouldn't exist without. More formally, an essential property F of x is such that x is F in every metaphysically possible world in which x exists, and in the closest possible or impossible worlds where F is not instantiated, x fails to exist.

Notice the difference between saying 'x is essentially F' and saying 'F is essential to x'. Your team's success is not essentially such that your website be advertised, but rather the advertising is essential to your team's success. Mafia Mike's protection is essential to Joey's existence, but Joey is not essentially protected by Mike. Doctors are essential to my existence, but I am not essentially a doctor. Whenever x is essentially F, F is essential to x; but sometimes F is essential to x without x being essentially F. This is just what our proposed analysis predicts. For 'a is necessarily F, and if there were no Fs, a wouldn't exist' obviously entails 'if there were no Fs, a wouldn't exist'.

In conclusion, if one recognizes non-trivial counterpossibles and distinguishes 'is essentially F' from 'F is essential to', one can offer a perfectly general modal analysis of 'essence'—an analysis that captures the philosophical sense, and entails the ordinary sense of the term.

3 Against non-trivial counterpossibles

Lewis offers intuitive justification for the vacuous reading of counterpossibles. He writes,

> Confronted by an antecedent that is not really an entertainable supposition, one may react by saying, with a shrug: If that were so, anything you like would be true! Further, it seems that a counterfactual in which the antecedent logically implies the consequent ought always to be true; and one sort of impossible antecedent, a self-contradictory one, logically implies any consequent. (1973, p. 24)

There are two justifications offered here. First, an antecedent that is not really an entertainable supposition invokes triviality. That may be right. But that by itself is not an objection to the modest proposal that counterpossibles are sometimes non-vacuously true/false. If a supposition is not really entertainable and therby invokes triviality, then at best that shows there are some trivial counterpossibles. It does not show that there are no non-trivial counterpossibles. To show that, one must presuppose that all counterpossibles involve antecedents that are not really entertainable. Part of what we have been arguing is that this is not the case—i.e., that counterpossibles sometimes express non-trivial consequences of very interesting entertainable, although impossible, positions.

The second justification articulated in the Lewis quote states that a counterfactual whose antecedent logically implies the consequent ought always to be true. That is, it ought to be that logically strict implication entails counterfactual implication:

$$\frac{P \mid\!- \ Q}{P \mathbin{\Box\!\!\rightarrow} Q}$$

The intuition here is of course theory-laden. Anyone who is convinced that there are some false counterpossibles will not share the intuition. That is because false counterpossibles are the counterexample to this inference. Indeed, strict implication implies counterfactual implication just in case all counterpossibles are vacuously true. One's theoretic intuitions about the implication between strict and counterfactual conditionals turn on one's intuitions about whether there are false counterpossibles. Since the above time slice of Lewis doesn't believe there are entertainable impossibilities (and furthermore since he favors the vacuously true, rather than the vacuously false, theory of counterpossibles), he doesn't think there are any false counterpossibles. Like his first justification for vacuism, Lewis' second justification may be viewed as an expression of skepticism about the entertainability of impossibility. As we have been arguing, this skepticism is not generally shared in ordinary and philosophical practice. Some impossibilities (indeed some impossibilities known to be impossible) are entertainable suppositions.[10]

Timothy Williamson objects to non-vacuism (Hempel Lectures 2006, and Williamson 2007). He asks us to consider someone who answered '11' to 'What is $5 + 7$?' but who mistakenly believes that he answered '13'. For the non-vacuist, presumably, (1) is false and (2) is true:

(1) If $5 + 7$ were 13, x would have got that sum right
(2) If $5 + 7$ were 13, x would have got that sum wrong

Williamson is not persuaded by the initial intuitiveness of such examples:

> ...they tend to fall apart when thought through. For example, if $5 + 7$ were 13 then $5 + 6$ would be 12, and so (by another eleven steps) 0 would be 1, so if the number of right answers I gave were 0, the number of right answers I gave would be 1.

Note Williamson's conclusion:

(3) If the number of right answers I gave were 0, then the number of right answers I gave would be 1.

The implicit reductio must be this. If (3) is true, then (1) and (2) are both true—contrary to what the non-vacuist supposes. For if I gave 0 right answers (in close worlds where $0 = 1$), then I also gave 1 right answer (in those worlds). Hence, I got that sum both right and wrong.

Williamson's abbreviated eleven-plus-one steps must be these:

(i) If $5 + 7$ were 13, then $5 + 6$ would be 12
(ii) If $5 + 7$ were 13, then $5 + 5$ would be 11
...
(xi) If $5 + 7$ were 13, then $5 + -4$ would be 2.
(xii) If $5 + 7$ were 13, then $5 + -5$ would be 1.

Despite appearances, getting to (3) from there is not a basic exercise. Williamson's reasoning seems to be that any world where $5 + -5 = 1$ is one where $0 = 1$, substituting '0' for '$5 + -5$'. Hence,

[10] More recently Lewis (2001, 176–177) has expressed sympathy with the idea that we seriously and non-trivial reason about impossibilities.

(xiii) If $5 + 7$ were 13 then 0 would be 1.

If $5 + 7$ were 13 (and I gave 0 right answers), then (since 0 would be 1) I would have given 1 right answer.

If this is Williamson's argument, then it's unsuccessful. First, substituting '0' for '$5 + -5$' is illicit, since as Williamson himself notes the non-vacuous counterfactual is hyperintensional. Hyperintensional operators do not permit substitutions of co-referring terms *salva veritate*.[11]

A second not unrelated problem for Williamson's position emerges in steps (i) through (xiii). These sub-conclusions hold, if the game is to evaluate the consequent of each at *deductively closed* worlds where $5 + 7 = 13$. But if there are non-trivial counterpossibles, the relevant worlds of evaluation must not be deductively closed—lest they collapse into the trivial world where everything is true. But once we deny deductive closure, each of Williamson's intermediary steps of reasoning, i.e., steps ii through xii, is in need of further justification.

Put it another way. Let the following world, W, fail to be deductively closed:

(W): {$5 + 7 = 13$, the number of right answers I gave wasn't 1, the number of right answers I gave was 0,...}

In contexts where W-worlds are closest, (2) is true and (1) false, as the non-vacuist predicts. For Williamson's argument to succeed, however, the relevant impossible worlds in which I gave 0 right answers *and* I gave 1 right answer must be closer than the relevant impossible W-worlds. This hasn't been shown. Indeed, pending further discussion, W seems closer to the actual world than Williamson's impossible world, since W conflicts with fewer salient background conditions.

Interestingly enough, Williamson on occasion sounds like a non-vacuist. Indeed his reasoning in places tacitly assumes there are non-trivial counterpossibles. For instance, while entertaining the view that all counterpossibles are vacuously false, he writes,

> If all counterpossibles were false, $\Diamond$A would be equivalent to A $\Box\!\!\rightarrow$ A, for the latter would still be true whenever A was possible; correspondingly, $\Box$A would be equivalent to the dual $\neg(\neg$ A $\Box\!\!\rightarrow \neg$ A). (Hempel Lecture, 2006)

The utterance is a counterpossible. If all counterpossibles are in fact true and necessarily so, as Williamson supposes, then what Williamson just said is vacuously true. But so is the opposite counterfactual—viz., 'If all counterpossibles were false, $\Diamond$A would *not* be equivalent to A $\Box\!\!\rightarrow$ A...'

So why didn't Williamson make his point with the latter counterfactual (i.e., the opposite), which says as much as the former given vacuism? Because the latter claim

[11] As Kallestrup (2009) argues, counterpossibles sometimes allow for substitution. For example, it seems plausible that 'Clark Kent' can be substituted for 'Superman' in 'if Superman had had the same parents as I do, then I would have had a brother'. However, we do not think this problem is insurmountable. Hyperintensional contexts do not always resist substitution even if they sometimes do. Of course, we need to give a principled account of when counterfactuals create opaque contexts. They create opaque context when the antecedent or consequent which result from substituting one term for another does not follow a priori from the original. Since we are likely to use 'Clark Kent' and 'Superman' in such a way as to pick out the same individual, 'Clark Kent has the same parents as I do' is an a priori implication of 'Superman has the same parents as I do'. So, substitution is legitimate.

is true, only if both claims are vacuous and *not* informative! Williamson did not intend to assert something vacuous and uninformative. Rather he tacitly assumes that the consequent contributes to the truth-value, and our understanding, of his claim. The vacuity treatment of counterpossibles that he favors is then not the view he intends his readers to employ when evaluating his assertion.

Since Williamson (above) and many others make deep and interesting claims with their counterpossible assertions, a non-vacuous reading is called for. Counterpossible philosophy is sometimes non-trivial and informative. The primary lesson of the section, however, is that the arguments against non-vacuism that we have looked at are all deficient in some way.

4 Impossible worlds

Since impossibilities entail everything in classical logic and we wish to avoid triviality, an impossible-worlds account of counterpossibles would seem to require us to weaken our consequence relation in some uniform manner. Quite naturally one might treat counterpossibles against a backdrop of paraconsistent logic, since paraconsistent logics reject ex falso quodlibet, i.e., the classical principle which allows us to derive an arbitrary proposition from a contradiction.[12] The key point would be that all paraconsistent consequences, but not all classical consequences, of the antecedent are counterfactually implied by the antecedent.

Daniel Nolan points out that a uniform weakening of the consequence relation is a bad idea, because no weakening can handle every impossibility that we might want to reason about. Here is a way to argue that very general point in terms of counterpossibles. Let L be the preferred/correct but weakened logic. Then, absurdly, the following counterpossible is not false:

If L were not the correct logic, then all and only L-theorems would be valid.

The conditional would not be treated as false, because ex hypothesi every world would be an L-world. Hence, there would be no worlds where the consequent is false, and so, no closest antecedent worlds where the consequent is false. Of course, L is an arbitrary logic. So any uniform weakening of the consequence relation would fail in the same way. By itself the strategy fails to capture all reasonable intuitions about the truth-values of counterpossibles.[13]

Nolan proposes an alternative. Countenance impossible worlds and treat them as logically anomalous. Nolan thinks Lewis' similarity-in-relevant-respects account can be extended to impossible worlds along the lines suggested above. He doesn't commit one way or another to what the similar-in-relevant-respects talk amounts to. But the following account is one possible way to go based on Nolan's discussion. With the exception of the absurd world—the world in which every proposition is

[12] See, for instance, Priest and Routley (1984) Insert Bib ref.

[13] The approach here is not to be thought of as combined with Nolan's approach of denying the deductive closure of impossible worlds. Once we deny the deductive closure of impossible worlds, we no longer have a need for uniformly weakening the consequence relation.

true—impossible worlds are maximal, but not deductively closed, sets of propositions or sentences.[14] The absurd world is both maximal and deductively closed. A set S is *maximal* iff for any proposition (sentence) p, either p or $\sim p$ (or both) is a member of S.[15] A set S is *not deductively closed* iff not every consequence of the propositions (sentences) in S is a member of S. Accordingly, p is true at an impossible world w iff p is a member of w.

Impossible worlds are then ordered relative to how relevantly similar they are to the actual world, given certain background facts. As Nolan puts it,

> some impossible worlds are more similar, in relevant respects, to our actual world than others. The "explosion" world—the impossible world where every proposition is true—is very dissimilar from our own. Indeed it seems to be one of the most absurd situations conceivable. On the other hand, the world which is otherwise exactly like ours, except that Hobbes succeeded in his ambition in squaring the circle (but kept it a secret), is far less dissimilar (1997, p. 544).

The metric is very much in the spirit of Lewis' metric for closeness of possible worlds. The impossible worlds with widespread violation of actual laws are farther from the actual world than are worlds with fewer violations.

5 Problems for impossible worlds

Despite its many virtues, the Nolanesque impossible worlds account of counterpossibles has a serious drawback. It yields the wrong results regardless of whether impossible worlds are treated as sets of Russellian propositions or sets of sentences (or sentence-like propositions).

Suppose worlds are sets of Russellian propositions. Since Russellian propositions contain the referents of terms as constituents, the proposition that water is not H_2O just is the proposition that H_2O is not H_2O. So, the following subjunctive comes out true:

(11) If water had not been H_2O, then H_2O would not have been H_2O.

But intuitively, if water had been something other than H_2O, then H_2O would still have been H_2O. We get the same results, mutatis mutandis, regarding terms such as '2' and '{{{}}}' which denote abstract entities.

The lesson is this: Unlike metaphysical necessity, counterpossible implication is a hyperintensional operator. It creates opaque contexts, and for this reason, the assumption that it operates on Russellian propositions is highly problematic.

Suppose instead that worlds are maximal (but not deductively closed) sets of *sentences*. Then how do we define 'closeness'? Here is one suggestion inspired by Nolan's

[14] Nolan leaves open whether worlds are concrete or abstract. However, since inconsistent beings are more bizarre than inconsistent sets, we will treat worlds as abstract. Moreover, in the appendix Nolan seems to allow for non-maximal worlds. We will treat worlds as maximal. That will provide us with the simplicity of a two-valued semantics.

[15] To avoid a number of red herrings, we may suppose for the time being that these sets contain only a countable number of propositions/sentences.

remarks: for any two worlds w_1 and w_2, w_1 is closer to the actual world than w_2 iff a greater number of (the relevant) sentences in w_1 than in w_2 express, relative to the speaker's context, Russellian propositions that are true in the actual world.

However, this account of closeness runs into difficulties. Consider (Model 1), in which two worlds w_1 and $w_2$that are exactly similar to each other and to the actual world, except that w_1 contains the sentences 'water is XYZ' and 'water is not H_2O', whereas w_2 contains the sentence 'water is H_2O' and 'water is not H_2O':[16]

w_1	w_2	
water is XYZ	water is H_2O	(Model 1)
water is not H_2O	water is not H_2O	

Since w_2 expresses a greater number of actually true Russellian propositions than w_1, w_2 is closer to the actual world than w_1. So, the subjunctive 'if water had not been H_2O, water would be H_2O' comes out true, contrary to intuition.

Of course, one might insist—with Nolan—that we need to minimize *formal* logical contradictions. For example, it might be required that for any two worlds w_1 and w_2, w_1 is closer to the actual world than w_2 iff w_1 expresses a greater number of actually true Russellian propositions than w_2, and w_1 does not contain a greater number of formal contradictions than w_2.

But this amendment is unlikely to help. For consider (Model 2), which is exactly like (Model 1) with the following changes. w_2 contains the sentence 'water is a monkey' (and all sentences that are the result of affixing an even number of negations to this sentence) and does not contain 'water is H_2O' (or its evenly negated counterparts):

w_1	w_2	
water is XYZ	water is a monkey	(Model 2)
water is not H_2O	water is not H_2O	

Since these worlds are formally consistent and contain the same number of relevant actual falsehoods, they are equally close. So, the subjunctive

(13) (Even) if water had not been H_2O, it would not have been a monkey

comes out false. For the antecedent is true at w_1 and w_2, but the consequent is false at w_1. However, (13) is true.

Consider another countermodel, (Model 3), in which w_1 and w_1 are exactly similar to each other and to the actual world, with the following exceptions: w_1 contains the sentences 'paraconsistent logic is correct' (and the sentences that are the result of affixing an even number of negations to this sentence) instead of the actually true sentence 'it is not the case that paraconsistent logic is correct' (and the sentences resulting from affixing an even number of negations to this sentence).[17] w_2 contains the sentences

16 It may be objected that that there are worlds even closer to the actual world than w_1 and w_2, for instance, a world in which water is not H_2O, not XYZ, not a monkey, not.... But clearly a non-vacuist ought to let context rule out such worlds; otherwise intuitively true counterfactuals such as 'if water had not been H_2O, it might have been XYZ' come out false.

17 Suppose with us that paraconsistent logic is not correct.

'paraconsistent logic is correct' and 'ex falso quodlibet is invalid' (and all sentences that are the result of affixing an even number of negations to these sentences) instead of the actually true sentences 'it is not the case that paraconsistent logic is correct' and 'it is not the case that ex falso quodlibet is invalid' (and the sentences that are the result of affixing an even number of negations to these sentences):

w_1	w_2	
paraconsistent logic is correct	paraconsistent logic is correct	(Model 3)
ex falso quodlibet is valid	ex falso quodlibet is not valid	

It should be noted that w_1 does not contain every sentence, for instance, it does not contain the sentence 'ex falso quodlibet is not valid'. w_1 is not deductively closed!

Now, w_1 and w_2 are formally consistent. Neither world contains both a sentence and its negation. But, since w_1 contains a greater number of relevant actually true sentences than w_2, w_1 is closer to the actual world than w_2. So, the following subjunctive absurdly comes out true:

(13) If paraconsistent logic were correct, then ex falso quodlibet would still be valid.

The immediate problem with the sentence treatment is that the corresponding closeness relation does not give weight to a priori connections between the sentences that constitute the worlds. For speakers who have 'paraconsistent logic' in their vocabulary, 'paraconsistent logic is correct' a priori implies 'ex falso quodlibet is not valid'. Likewise, for speakers who have 'water' in their vocabulary, 'water is not H_2O' a priori implies 'water is not a monkey', as 'water is not a monkey' is a priori. We will now develop an epistemic account of subjunctive conditionals that handles a priori connections.

6 An alternative epistemic account of counterpossibles

The account of subjunctives we prefer is a variation on the Nolanesque account considered above but one which treats the closeness relation as partially epistemic. In particular, we propose to make important use of the a priori. We follow epistemic two-dimensionalists in taking the notion of apriority as highly context-sensitive. It varies with one's use of the elements of the sentence in question (see Chalmers 1996, 2002, 2006, manuscript, Chalmers and Jackson 2001). Moreover, like the two-dimensionalists, we do not propose to give an account of apriority. The notion of apriority is taken as basic. But our notion will deviate from that of the two-dimensionalists in one important respect. Two-dimensionalists stipulate that all mathematical and logical truths are a priori. We do not make this stipulation. Intuitively, 'paraconsistent logic is correct' does not a priori imply 'ex falso quodlibet is valid' for every speaker in every context. But the a priori implication would obtain if all logical truths were a priori, and for the simple reason that, ex hypothesi, ex falso can be formulated as a logical truth. Our default assumption will be that no logical truth is a priori for a subject (in spite of the fact that any logical truth is necessary). To distinguish our notion from the standard one, let us call it 'a priori*'. 'A priori* implication' may be characterized as follows:

(A priori* Implication)

For a speaker s in a context c, P *a priori* implies* Q iff for s in c, Q is a relevant a priori consequence of P.

Let the totality of worlds be the maximal sets of sentences. The possible worlds will be the deductively closed worlds, with the exception of the absurd world (i.e., the world containing every sentence). The impossible worlds will be all worlds that are not deductively closed, and the absurd world.

Worlds are ordered in terms of comparative closeness. How do we define closeness? As mentioned earlier, the problem with the Nolan-inspired analysis of counterpossibles is that its closeness relation doesn't give any weight to a priori connections between the sentences. But it is not sufficient that the closeness relation gives weight to a priori connections between sentences while preserving formal consistency. In the spirit of Lewis' account, the closeness relation must give weight, first and foremost, to the background facts kept fixed in the conversational context. For example, if it is assumed that the path from the 10th floor to the ground is unobstructed, then if John were to jump out from the 10th floor, he would be hurt. If, on the other hand, it is assumed that John is a sensible guy who wouldn't jump from the 10th floor unless there was a safety net, then if John were to jump out from the 10th floor, he would not get hurt (Jackson 1987, p. 9). A priori* connections then may make a difference when impossible worlds are tied in terms of how compatible they are with the background facts.

To summarize: the closeness relation must, first of all, give weight to background facts and secondly, it must give weight to a priori* connections. We propose to define closeness as follows:

(Closeness)

1. Minimize discrepancies with backgrounds facts (fixed by context)
2. If needed, maximize a priori* connections

Or slightly less informal:

For any two impossible worlds w_1 and w_2, w_1 is closer to the base world than w_2 iff

(a) w_1 does not contain a greater number of sentences formally inconsistent with the relevant background facts (held fixed in the context) than w_2 does. And if w_1 and $w_2$contain the same number of sentences formally inconsistent with the relevant background facts (held fixed in the context):

(b) w_1 preserves a greater number of a priori* implications between sentences than w_2 does.

Background facts will include facts about laws of nature, facts about what the world is like, and so on (see Lewis 1979).

Our definition of closeness allows that impossible worlds tied in the relevant respects articulated by Nolan can vary, owing to a difference in what is a priori* for the speaker. Suppose it is a priori* for s that H_2O is not a monkey but that it is not a priori* for s that water is H_2O, perhaps because she hasn't yet discovered the nature of water. Consider again (Model 2). We suppose w_1 and w_2 are exactly similar to each other and to the actual world, except that w_1 contains the sentences 'water is not H_2O' and 'H_2O is a monkey', and w_2 contains the sentences 'water is not H_2O' and 'water is XYZ'.

w_1	w_2	
water is not H_2O	water is not H_2O	(Model 2)
water is a monkey	water is XYZ	

Since 'H_2O is not a monkey' is a priori* but 'water is H_2O' is not, it follows that w_2 is closer to the actual world than w_1, which is as it should be. Considerations like this preserve the intuition that water would not be monkey even if it were not H_2O.

Likewise, with Model 3. w_1 and w_1 are exactly similar to each other and to the actual world, with the following exceptions: w_1 contains the sentences 'paraconsistent logic is correct', and w_2 contains the sentences 'paraconsistent logic is correct' and 'ex falso quodlibet is not valid':

w_1	w_2	
paraconsistent logic is correct	paraconsistent logic is correct	(Model 3)
ex falso quodlibet is valid	ex falso quodlibet is not valid	

Given the notion of apriority assumed here, 'paraconsistent logic is correct' a priori* implies 'ex falso quodlibet is not valid' in normal contexts. Hence, w_2 is (at least in normal contexts) closer to the actual world than w_1. Again this is as it should be.

The two-dimensional account just offered yields the right results for other counterpossibles as well. 'H_2O is H_2O' is a priori* for all minimally rational speakers. So, worlds at which 'water is not H_2O' and 'H_2O is H_2O' are both true are closer to the actual world than worlds at which 'water is not H_2O', and 'H_2O is not H_2O' are true. So, the subjunctive 'if water were not H_2O, then H_2O would still be H_2O' comes out true.

The present account handles these cases nicely because it takes the truth-value of the relevant subjunctive conditionals to be determined in part by what is a priori* for the speaker in the conversational context. 'It is a priori* that' creates opaque contexts, as is illustrated by the fact that 'it is a priori* that water is water' and 'it is *not* a priori* that water is H_2O' may both be true. Since 'it is a priori* that' creates opaque contexts, and the truth-conditions for counterpossibles are determined in part by what is a priori* for the speaker, we have an explanation for the opacity of counterpossibles.

7 Objections

We now turn to some potential objections to the account just offered.

1. Williamson (2006, Chap. 5) offers a general objection to impossible worlds accounts of counterpossibles. The thrust of the objection is that once we countenance impossible worlds in our account of counterfactual conditionals, counterfactuals create opaque contexts where there should not be. Suppose, as we do, that the counterfactual conditional creates opaque contexts. Then the following clearly valid inference should be treated as invalid:

(Inference 2)

If the rocket had continued on that course, it would have hit Hesperus.
Hesperus = Phosphorus.
Therefore, if the rocket had continued on that course, it would have hit Phosphorus.

We agree with Williamson that this inference is valid. Moreover, our account does not require us to say otherwise. Yes, the counterfactual conditional is hyperintensional. However, we need not throw out the baby with the logically ill-behaved bath water. Our logical principles may be restricted accordingly. All the typical rules governing counterfactuals are valid, when the antecedent is possible. The above argument, by contrast, does not involve an impossible antecedent.

2. It might be thought that two-dimensional accounts of subjunctives yield the wrong results for subjunctives with 'actually' in the antecedent or consequent. Consider, for instance:

(14) If $2 + 2$ had been 5, $2 + 2$ would actually have been 5.

On the standard philosophical account of 'actually', 'actually, φ' is true at a world w iff φ is true at the actual world:

Rigidification-1: Aφ is true at a world w iff φ is true at the actual world

Given the philosophical account and that actually $2 + 2 = 4$, it might be argued that (14) ought to be false. However, it may seem that our epistemic account predicts that (14) is true. For the epistemic account holds that worlds that preserve more relevant a priori* connections are closer to the actual world. But '$2 + 2 = 5$' a priori* implies '$2 + 2$ is actually 5'. So, the closest world in which $2 + 2 = 5$ is one in which $2 + 2$ is actually 5.

Our reply is that the epistemic account does not make this predication, given ordinary contexts. Consider (Model 4):

w_1	w_2	
$2 + 2 = 5$	$2 + 2 = 5$	(Model 4)
$2 + 2$ is not actually 5	$2 + 2$ is actually 5	

If '$2 + 2$ is not actually 5' is among the background facts kept fixed in the context, then a world w_1 at which both $2 + 2 = 5$ and $2 + 2$ is not actually 5 contains a smaller number of sentences formally inconsistent with the background facts than a world, w_2, in which $2 + 2 = 5$ and $2 + 2$ is actually 5. And so w_1 is closer to the actual world than is w_2. So, if '$2 + 2$ is not actually 5' is among the background facts kept fixed in the context, then (14) comes out false.[18]

If, by contrast, '$2 + 2$ is not actually 5' is not among the background facts kept fixed in the context, then, given salient a priori connections, w_2 is closer to the actual world than w_1 is. So, if '$2 + 2$ is not actually 5' is not among the background facts

[18] A different line of reply would be to insist that 'actually' does not function univocally in ordinary language. For an argument for a thesis along those lines see Brogaard (2007).

kept fixed in the context, then (14) is intuitively true, and our account gives the right result.

As it turns out, subjunctive conditionals with ‘actually’ cause trouble for the standard impossible worlds account, but not for the epistemic account. Suppose you utter the following subjunctive on a sunny day:

(15) If it had actually been raining, it would not have been raining.

As it is not raining in the actual world, there is no possible world in which it is actually raining. In fact, if Rigidification-1 is true, there cannot be any worlds in which the antecedent of (15) is true (not even the absurd world). For if the antecedent of (15) is true at a world, then by Rigidification-1, it is raining in the actual world, but ex hypothesi, it is not.

Those who think that ordinary-language occurrences of ‘actually’ must be treated as rigidifiers may propose the following alternative to Rigidification:

Rigidification-2: $A\varphi$ is true at a *possible* world w iff φ is true at the actual world

But given Rigidification-2, we can evaluate (15) only if we stipulate that if the antecedent contains the word ‘actually’, and if the antecedent is impossible, then the subjunctive is vacuously true. But this is ad hoc. Moreover, if we make the stipulation, then the account will succumb to the difficulties it purports to overcome. For its principal motivation was to account for our intuitions about counterpossibles. Yet (15) surely seems false.

On the present account, (15) does not present a particularly difficult problem. ‘It has actually been raining’ a priori* implies ‘it has been raining’. So, the closest worlds at which it has actually been raining are worlds in which ‘it has been raining’ is true. So, (15) is predicted to be false.

3. There is, however, another, more intricate, problem for epistemic accounts of subjunctives, viz. that of accounting for philosophical intuitions about rigid terms embedded in subjunctive environments. Suppose Twin-Earth has become almost dry. Dying of thirst the few remaining Twin-Earthlings set out to find a lake located 10 miles away.[19] Sadly, they die from dehydration just before they reach their destination. Now, consider:

(16) If the inhabitants of Twin Earth (a planet at which the watery stuff that flows in rivers, lakes and oceans is XYZ) had reached the lake, they would have found water.

Philosophers judge (16) to be false. If the Twin-Earthlings had reached the lake, they would have found XYZ, not water. Yet if ‘water’ is used interchangeably with ‘watery stuff that flows in rivers, lakes and oceans’ (that is, if ‘water’ is a priori* equivalent to ‘watery stuff’), then ‘the inhabitants of Twin Earth reach the lake’ a priori* implies ‘the Twin Earthlings find water’. So, ‘the Twin Earthlings found water’ is true at all the closest worlds in which ‘the Twin Earthlings reach the lake’ is true.

However, this is not an objection to the present account. The reported intuitions are those of philosophers. In typical philosophical contexts, ‘water is H_2O, not XYZ’

[19] The story is based on Weatherson (2001) story about explorers in Twin Australia.

is a background fact. So, worlds in which 'the inhabitants of Twin Earth reached the lake' and 'the Twin Earthlings found water' are both true are in violation of the background facts. Worlds in which 'the inhabitants of Twin Earth reached the lake' and 'the Twin Earthlings did not find water' are not. So, (16) comes out false, and our account predicts this result.

Not unsurprisingly, untutored folks judge (16) to be true. This, of course, owes to the fact that 'water is H_2O, not XYZ' is not typically a background fact in ordinary contexts. Untutored folks are quite happy to judge as true statements such as 'the quality of your coffee will depend on the chemical composition of the water you use' and 'the collected samples of water contained 0.02 % lead'.

There is a similar bias concerning the truth-value of 'if water had been XYX and not H_2O, then water would (still) have been H_2O'. If it is assumed in the conversational context that 'water is H_2O, not XYZ' is true, then the closest worlds in which water is XYZ and not H_2O are worlds in which water is H_2O; hence, the subjunctive comes out true. If, on the other hand, it is not assumed in the conversational context that 'water is H_2O, not XYZ', then worlds exactly like the actual except that the drinkable liquid in oceans and lakes is XYZ and not H_2O are much closer than worlds in which water is both H_2O and not H_2O.

8 Conclusion

One of the prime objections to Lewis's account of subjunctive conditionals is that it yields counterintuitive results for subjunctives with necessarily false antecedents. Nolan has argued that Lewis' account can be extended to account for subjunctives with necessarily false antecedents if we allow impossible worlds to be ordered by similarity. On this account, impossible worlds are not deductively closed, and so, can resemble the actual world more or less closely.

Unfortunately, as we have seen, the account runs into trouble, regardless of whether he takes worlds to be sets of Russellian propositions or sets of sentences. This is because, on Nolan's account, closeness does not give weight to a priori connections.

We have argued that if we take account of a priori connections, the informativeness of subjunctive conditionals can be accounted for. A subjunctive of the form 'if p were the case, q would be the case' is true at a base world w iff 'q' is true at the 'p'-worlds closest to w. Here are two criteria for maximizing closeness. 1. Minimize discrepancies with relevant backgrounds facts (fixed by context). 2. If needed, maximize relevant a priori connections. On our account, logical truths are not relevantly a priori by default. The account, we believe, allows for a clear modal distinction between essence and de re necessity, and makes room more generally for the non-trivial counterpossibles we know and love.

Acknowledgments The paper has benefited from correspondence with Mike Almeida, Alan Baker, Johan van Benthem, J.C. Bjerring, John Broome, Ross Cameron, David Chalmers, Tom Crisp, Louis deRosset, Matti Eklund, Al Hajek, Carrie Jenkins, Jesper Kallestrup, Kathrin Koslicki, Stephanie Lewis, John MacFarlane, Mark Moyer, Daniel Nolan, Laurie Paul, Agustin Rayo, Michael Rea, Gillian Russell, Jonathan Schaffer, Wolfgang Schwartz, Ted Sider, Jim Stone, Michael Strevens, Weng Hong Tang, Ken Taylor, Robbie Williams and Tim Williamson, discussions with the participants in a reading group at Washington

University on Williamson's *The Philosophy of Philosophy* spring 2007 and audience discussions in Arizona, Copenhagen, Edinburgh, Paris, Sydney, and at ANU. Special thanks to David Chalmers who provoked us to think about counterpossibles.

References

Brogaard, B. (2007a). Comments on Delia Graff Fara's coincidence by another name. Arizona Ontology Conference, Jan 2007. [http://brogaardb.googlepages.com/CommentsonDeliaGraffFara.doc].
Brogaard, B. (2008). Sea-battle semantics. *Philosophical Quarterly 58*(231), 326–335
Brogaard, B., & Salerno, J. (2008). Counterfactuals and context. *Analysis 68*(1), 39–46
Chalmers, D. (1996). *The conscious mind: In search of a fundamental theory*. Oxford: Oxford University Press.
Chalmers, D. (1998). The tyranny of the subjunctive. [http://consc.net/papers/tyranny.htlm].
Chalmers, D. (2002). On sense and intension. *Philosophical Perspectives, 16*, 135–182.
Chalmers, D. (2006). The foundations of two-dimensional semantics. In M. Garcia-Carpintero & J. Macia (Eds.), *Two-dimensional semantics: Foundations and applications* (pp. 55–139). Oxford: Oxford University Press.
Chalmers, D. J. (2011). Propositions and attitude ascriptions: A fregean account. *Nous*, *45*, 595–639.
Chalmers, D., & Jackson, F. (2001). Conceptual analysis and reductive explanation. *Philosophical Review, 110*, 315–361.
Fine, K. (1994). Essence and modality. *Philosophical Perspectives, 8*, 1–16.
Hájek, A. Most counterfactuals are false, Manuscript.
Jackson, F. (1987). *Conditionals*. Oxford: Blackwell.
Kallestrup, J. (2009). Conceivability, rigidity and counterpossibles. *Synthese*, *171*(3), 377–386.
Kim, S., & Maslen, C. (2006). Counterfactuals as short stories. *Philosophical Studies, 129*, 81–117.
Kment, B. (2006a). Counterfactuals and explanation. *Mind, 115*, 261–310.
Kment, B. (2006b). Counterfactuals and the analysis of necessity. *Philosophical Perspectives, 20*, 237–302.
Kripke, S. (1980). *Naming and necessity*. Cambridge, MA: Harvard University Press.
Lewis, D. (1973). *Counterfactuals*. Oxford: Blackwell.
Lewis, D. (1979). Counterfactual dependence and times arrow. *Noûs, 13*(4), 455–476.
Lewis, D. (1981). Ordering semantics and premise semantics for counterfactuals. *Journal of Philosophical Logic, 10*(2), 217–234
Lewis, D. (1986). *On the plurality of worlds*. Oxford: Basil Blackwell.
Moore, G.E. (1919). *External and internal relations*. Aristotelian Society.
Nolan, D. (1997). Impossible worlds: A modest approach. *Notre Dame Journal for Formal Logic, 38*, 325–527.
Priest, G., & Routley, R. (1984). Introduction: Paraconsistent logics. *Studia Logica, 43*(1–2), 3–16.
Wierenga, E. (1998). Theism and counterpossibles. *Philosophical studies, 89*(1), 87–103.
Vander Laan, D. (2004). Counterpossibles and similarity. In F. Jackson & G. Priest (Eds.), *Lewisian themes*. Oxford: Oxford University Press.
Weatherson, B. (2001). Indicatives and subjunctives. *Philosophical Quarterly, 51*, 200–216.
Williamson, T. (2007). *The Philosophy of Philosophy*, Blackwell.