Rieger, A. (2017) Was Quine right about subjunctive conditionals? *Monist*, 100(2), pp. 180-193. (doi:10.1093/monist/onx003)

This is the author's final accepted version.

There may be differences between this version and the published version. You are advised to consult the publisher's version if you wish to cite from it.

http://eprints.gla.ac.uk/133977/

Deposited on: 11 January 2017

Was Quine right about subjunctive conditionals?

Abstract

Given his hostility to intensional locutions, it is not surprising that Quine was suspicious of the subjunctive conditional. Although he admitted its usefulness as a heuristic device, in order to introduce dispositional terms, he held that it had no place in a finished scientific theory. In this paper I argue in support of something like Quine's position. Many contemporary philosophers are unreflectively realist about subjunctives, regarding them as having objective truth values. I contest this. "Moderate realist" theorists, such as Lewis and Stalnaker, admit that subjunctives are context-relative and often indeterminate; I argue, using some examples from the contemporary literature on conditionals, that these features are deeper and more widespread than they think. "Ultra realist" theories, which deny any indeterminacy, are not credible. Hence subjunctives are unsuitable for certain purposes, in particular the description of mind-independent reality.

A boy is going on his first date. Worried that conversation will dry up, he asks his father for advice. "Just stick to the three safe topics, son: food, family and philosophy." So the boy goes on the date. Conversation flags

*but the boy remembers his father's advice. "Do you like potato pancakes?"
he asks. "No", the girl replies. There is an uncomfortable silence. "Do you
have a sister?" asks the boy. "No", the girl replies again. Another
uncomfortable silence. So in desperation he tries philosophy: "If you had a
sister, would she like potato pancakes?"*

Very old joke[1]

1 Introduction: Quine on conditionals

Quine was no great friend of subjunctive[2] conditionals. Given Quine's hostility to non-extensional locutions, this is not surprising.[3] Moreover, Quine gives specific examples to motivate pessimism as to the prospects of developing an account of them which would reveal them to be, after all, suitable for scientific discourse. In *Methods of Logic*, Quine writes:

---

[1]   A version of the joke ("Well, if you had a brother do you think he'd like cheese?") appears (Act II, Scene 2) in the play *Our American Cousin*, written in 1858 and famous as the play Lincoln was watching when he was assassinated.

[2]   As is standard, I use 'subjunctive' and 'counterfactual' interchangeably for the class of conditionals under discussion, though as is well-known neither term is ideal, since there are specimens which are not in the subjunctive mood as well as ones with true antecedents.

[3]   One might have expected Quine to recommend the impeccably extensional material conditional as an account of the *indicative* conditional. Somewhat surprisingly, however, he seems to favour (eg Quine 1960, 226) an account according to which such a conditional has a truth-value gap when the antecedent is false. On the other hand, he is quite explicit (1982, 22, 95) that statements of the form "All Fs are G" can be analysed perfectly as universally quantified material conditionals. The position is uncomfortable: "All Fs are G" seems to entail all instances of "If *a* is an F, *a* is G", but it will not on Quine's account if *a* is not an F. I have elsewhere (Rieger 2006, 2012, 2013) defended the viability of the material account, and will not discuss it further here.

It may be wondered, indeed, whether any really coherent theory of the contrafactual conditional is possible at all, particularly when trying to adjudicate between such examples as these:

If Bizet and Verdi had been compatriots, Bizet would have been Italian;

If Bizet and Verdi had been compatriots, Verdi would have been French.

(Quine 1982, 23)

In *Word and Object*, Quine gives another pair of examples:

The subjunctive conditional depends, like indirect quotation and more so, on a dramatic projection: we feign belief in the antecedent and see how convincing we then find the consequent. What traits of the real world to suppose preserved in the feigned world of the contrary-to-fact antecedent can be guessed only from a sympathetic sense of the fabulist's purpose in spinning his fable. Thus consider the pair (Goodman's, nearly enough)

If Caesar were in command, he would use the atom bomb;

If Caesar were in command, he would use catapults.

We are likelier to hear the former, but only because that one is likelier to fit a lesson that a speaker would try to dramatize...the subjunctive conditional is an idiom for which we cannot hope to find a satisfactory general substitute in realistic terms… the subjunctive conditional has no place in an austere canonical notation for science…

(Quine 1960, 222, 225)

Quine does allow for the scientific use of dispositional terms, such as *fragile*, introduced by a subjunctive conditional – thus, we call an object fragile if it would break if it were struck.[4] But "each disposition, in my view, is a physical state or mechanism" (1974, 10). Dispositional terms are merely convenient place-holders, ultimately to be replaced by categorical terms – molecular structure, or whatever – as science progresses. Does this

---

[4]   More recent work has shown considerable problems with the simple analysis of dispositions in terms of subjunctives which Quine endorses, for example the issue of *finkishness*. These complications need not concern us here.

mean that dispositional terms (and subjunctive conditionals) are eliminable? No. We can expect that each such term will, in time, be replaced by the appropriate categorical basis. But, since science is always in a state of development, we cannot expect there to be a time at which all such replacements will have occurred. (To ignore this distinction would be to commit a quantifier-shift fallacy.) Thus dispositional terms (and subjunctives) will always be with us, although the conditionals cannot be given a uniform analysis which would determine their truth conditions. Though they have an invaluable heuristic use, they are not, ultimately, scientifically respectable.[5]

Few would now endorse Quine's pessimism. A huge amount of theoretical work has been done on subjunctive conditionals, most famously by Lewis, who not only gave a theory of them but put them to work extensively, for example in his theory of causation. Counterfactuals are part of the daily discourse of philosophers.

But are they really in good standing? I suggest not. The difficulties Quine raises have really been swept under the carpet, not solved. And some of the extensive work on conditionals that has taken place in recent decades has revealed new problems.

In the rest of this paper I shall proceed as follows. I shall first review the mainstream theories of counterfactuals, those of Lewis and Stalnaker. Both philosophers advance only a modest realist about their accounts, and admit openly that counterfactuals display context-sensitivity, vagueness and indeterminacy. They both think, however, that these can be overcome enough to justify their use. If this is so, I argue, one should minimally expect to find a broad consensus on our intuitions concerning counterfactuals and some basic principles governing their logic. But, on the contrary, the fundamentals are in disarray, as I shall show by considering some problems discussed in the recent literature. Finally I shall briefly consider, but reject, some versions of realism more robust than those of Lewis and Stalnaker.

2. Lewis

The most influential writer on counterfactuals been Lewis. His possible worlds analysis of their truth conditions, first given in (1973), has become close to orthodoxy in philosophy. Famously, Lewis is an extreme realist about possible worlds: for him they are as

---

5   Quine's views on dispositions are set out in most detail in *The Roots of Reference* (1974, §3).

concrete as the actual world. Though few have followed him in this, the essential idea – roughly,[6] that A > C[7] is true iff in all the most similar[8] worlds to the actual world where A is true, C is true – has become part of most philosophers' intellectual toolkit.

What, though, of the comparative similarity relation? One might expect Lewis, in line with his realism, to hold that there are objective facts about similarity, to which we have some (though no doubt fallible) access. This would give counterfactuals a firm foundation. But he writes as follows:

> Overall similarity consists of innumerable similarities and differences in innumerable respects of comparison, balanced against each other according to the relative importances *we attach*[9] to those respects of comparison. Insofar as these relative importances differ from one person to another, or differ from one occasion to another, or are indeterminate even for a single person on a single occasion, so far is comparative similarity indeterminate.
>
> (1973, 91)

> He quotes with approval Nelson Goodman: "Importance is a highly volatile matter, varying with every shift of context and interest, and *quite incapable of supporting the fixed distinctions that philosophers seek to rest upon it*.[10]"
>
> (Lewis 1973, 92; the quotation is from Goodman 1970, 27)

---

6   Only roughly, because there may be no collection of possible worlds which is maximally similar to the actual world. The subtlety is not important for current purposes.

7   I use the symbol '>' to stand for the subjunctive conditional.

8   Actually, it is somewhat misleading to think of a single comparative similarity relation defined on the universe of possible worlds, as this imports structure not present in Lewis's official semantics, for example, symmetry. But as Lewis himself points out (1973, 51), world *j* may look similar from the point of view of world *i*, yet *i* not look similar from the point of view of world *j*, since different things are important at each world. That is, the similarity relation is liable to vary from world to world. Few have noticed this and I ignore it here: see Goodman (2015a).

9   My italics.

10   My italics.

Lewis says that the vagueness surrounding the comparative similarity relation is "the same sort of vagueness that arises if I say that Seattle resembles San Francisco more closely than it resembles Los Angeles" (92). The answer depends on whether we attach more importance to architecture, climate, politics, and so on. The case of possible worlds is likely to be, if anything, worse, since they are bigger and more complex than cities. And the quotations from Lewis make clear that the problem is not simply one of context-sensitivity of the sort involved when "I am now hungry" has to have a time and person supplied in order to have a determinate truth condition. It depends on the speaker's intention, which may remain impenetrable however much is known about the context of utterance, and may indeed itself be indeterminate (perhaps the speaker has nothing in particular in mind).

All this, one might think, is grist to Quine's mill. How we fill in what Quine above called the "traits of the real world to suppose preserved in the feigned world of the contrary-to-fact antecedent" is, on Lewis's account, *at best* a matter of intersubjective agreement, saturated in the particular local interests of humans. At worst it is entirely indeterminate. Either way, it seems too flimsy for counterfactuals to carry much weight in either science or philosophy.

Lewis, however, thinks not. His idea is that "the relative importances of respects of comparison, and thereby the comparative similarity of worlds, are at least roughly fixed" (93). His approach is essentially supervaluational: he admits that there are many ways to make precise the vague similarity relation, but contends that in many cases, counterfactuals will have a determinate truth value in all "permissible" ways to do this, that is, ways consistent with context and normal usage. He concedes, though, that "some sensitive counterfactuals are so vague as to be unsuitable for use in serious discourse" (94).

What reason does Lewis have for his optimism that, in a good proportion of cases, there will be enough intersubjective agreement that counterfactuals have determinate truth values, despite their sensitivity to how the vagueness of the similarity relation is resolved? The best evidence is surely that we do, in fact, often agree about their truth values. But how much can rest on this observation?

For one thing, our intuitions about counterfactuals turn out to be quite confused. I discuss below some examples from the literature which show up problems. If we cannot agree, for example, about fundamental logical principles concerning counterfactuals, their suitability for scientific discourse is in doubt. Secondly, even if we were in complete agreement, would that be enough? The agreement could be the result of systematic confusion and prejudice, and there is no obvious way to test them. Unless we have some account of their epistemology which gives some reason to think we are tracking genuine truth values, we are on flimsy ground here. And the epistemology is notoriously hard to

provide, and only made worse by the introduction of possible worlds à la Lewis, since we are entirely isolated from them.

In (1973) Lewis said little about how we make our judgments of similarity. This led to some criticisms: for example in a famous review Fine (1975) suggested that "If Nixon had pressed the button, there would have been a nuclear holocaust" wrongly comes out false on Lewis's semantics, since a world in which, for example, the wiring of the button fails is more similar to the actual world than one in which the system functions as designed and the missiles are launched. In response, Lewis (1979) gave his well-known recipe for determining similarity, justifying it by reference to the Nixon example:

(1)   It is of the first importance to avoid big, widespread, diverse violations of law.

(2)   It is of the second importance to maximize the spatio-temporal region throughout which perfect match of particular fact prevails.

(3)   It is of the third importance to avoid even small, localized, simple violations of law.

(4)   It is of little or no importance to secure approximate similarity of particular fact, even in matters that concern us greatly.

(Lewis 1979, 472)

There is a great deal of literature on whether these are correct, how they can be improved, and indeed whether any such approach can work (see eg Tichy (1976), Edgington (1995, 257) and Bennett (2003, *passim*)). Here I just want to consider the *status* of these rules. Lewis is quite explicit that they are "reverse engineered" to fit our pre-existing intuitions about counterfactuals; he makes no attempt to give an independent argument in their favour:

...we must use what we know about counterfactuals to find out about the appropriate similarity relation – not the other way around.

(467)

He is also explicit that the similarity relation given by the four rules is only a "default" one and that different ways of resolving the vagueness in the similarity relation are appropriate in different contexts (457).

Does Lewis's theory show that Quine's negative view of counterfactuals is mistaken? I think not. Grant for the sake of argument that the theory is at least a reasonably successful account of our intuitions about counterfactuals, perhaps better than that. Nevertheless, in the absence of some account as to why our intuitions are a reliable guide to

an independently existing *truth* about these conditionals, which seems particularly unlikely to be forthcoming given Lewis's emphasis on the multiple ways to resolve the vague similarity relation, what we seem to have been given is a piece of conceptual analysis in a narrow sense, of a concept saturated in interests local at least to the human race and likely particular societies and times. It is perhaps analogous to accounts in applied ethics of concepts such as *guilt* and *envy*. This is not to say such analyses are not valuable. They may even be of considerable importance (for example in legal discussion). But in carrying out such analyses we do not take ourselves to be uncovering fundamental truths of the universe, in the manner of a Quinean mature science, or naturalistic philosophy.

This will carry over to any application of counterfactuals in analysing other notions. For example, Lewis produced a well-known account of causation in terms of counterfactual dependence. He is quite open that "the vagueness of similarity does infect causation, and no correct analysis can deny it" (1974, 560). But it is my impression that most working on counterfactual theories of causation take themselves to be exploring something that is more objective and real than Lewis's foundations seem to allow.

3 Stalnaker

Another important theorist in the area is Robert Stalnaker. In a series of papers starting with (1968), he developed a theory similar in outline to Lewis's. Salient differences are the absence of Lewis's extreme realism about possible worlds; that A > C is true if C is true at the (single) closest world where A is true, rather than *all* the closest worlds as in Lewis's theory; and that Stalnaker does not attempt to give a Lewis-style account of which the similarity relation, instead leaving his "selection function" which picks out the nearest world for a given world and context as unanalysed.

Stalknaker does not take his semantic theory to be a full account of the truth conditions of conditionals, because it leaves open how the selection function is to behave. Somehow it depends on the context of utterance, the speaker's intentions and the speaker's linguistic community. As already noted, he does not attempt to give details. But he does make a number of claims about it. He devotes a chapter of (1984) to the question of whether his theory is realist about counterfactuals, arguing that it is a "modest realist" one. Like Lewis, his view is that at least sometimes, counterfactuals are determinately true or false, although "in application there is great potential for indeterminacy in the truth conditions for counterfactuals" (137). His view of the Quine Bizet-Verdi pair is that both conditionals are indeterminate. The selection function fails to pick out a unique world, as

there is a tie for similarity between worlds where both are Italian and both are French. (In contrast, in Lewis's semantics both conditionals are false. I discuss this further below.)

Overall, therefore, Stalnaker's position is similar to Lewis's in terms of his views on the determinacy of counterfactuals. His chapter includes a brief discussion (147-153) of whether counterfactuals are of use in science, replying to writings by Mackie (1973) and van Fraassen (1980) suggesting that they are not. Mackie argues that possible worlds approaches fail to provide truth conditions. He gives the example "If you had struck that match it would not have lit, said of a wet match".   He remarks that the conditional

> ...will be true if the closest possible world in which the match is struck is one where it is struck while still wet, our causal laws being still in force. But it is false if the closest possible world is one in which the striker, being knowledgeable and observant, would not strike a wet match, but takes care to dry this one before striking it.

> (89)

All we have been given, Mackie complains, is acceptability conditions, which depend on which worlds a speaker regards as closer – which seems little different from saying that it depends on which features of the world the speaker decides to retain in considering a situation where the antecedent is true.

Van Fraassen gives (1980, 116) an example similar in structure and draws a similar conclusion. Whilst not claiming counterfactuals do not have truth conditions at all, he regards their context-sensitivity as making them unsuitable for science:

> The truth-value of a conditional depends in part on the context. Science does not imply that the context is one way or another. Therefore science does not imply the truth of any counterfactual – except in the limiting case of a conditional with the same truth-value in all contexts.

> (1980, 118)

In reply to Mackie, Stalnaker points out that although the counterfactual *sentence* may not have been given determinate truth conditions, this does not mean that a counterfactual *proposition* cannot have them. The context plays the role of determining which proposition is picked out by the sentence. And against van Fraassen, he questions whether science is really as context-free as van Fraassen assumes. Do not domains of quantification, for example, vary according to context?

It might be too strong to claim that *any* context dependence rules out a statement from being a legitimate part of science. But, in the case of an indexical or an implicit domain of quantification, it will typically be possible to rephrase the original statement in a way that

is not-context dependent. On the other hand since the range of what is held fixed and what is allowed to vary can be arbitrarily complicated, this will not in general be possible with counterfactuals. Like Lewis, Stalnaker has failed to make a convincing case that the context-dependency of counterfactuals does not rule them out of a respectable science.

Lewis and Stalnaker are not the only philosophers to have produced theories of counterfactuals, though they have been the most influential. I shall not discuss any other theories in detail, but whatever the theory, similar questions will arise: what is kept fixed, and what allowed to vary, when the conditional is evaluated?

4 Morgenbesser's coin

One seeking to argue with Lewis and Stalnaker that the admitted vagueness and context-sensitivity of counterfactuals does not render them useless for serious purposes will surely welcome as a datum widespread agreement as to their truth values. In the next sections I will discuss some problem cases in the recent literature on conditionals which suggests that on some fundamental matters this agreement will be hard to obtain.

Slote (1978, fn 33) gives the following example, which he attributes to Sidney Morgenbesser. A friend offers you good odds against heads coming up on the toss of an indeterministic coin. You refuse the bet, the coin is tossed, and it does come up heads. The problem concerns the counterfactual

(M)   If you had bet on heads, you would have won.

A natural reaction is that (M) is true. One can then test any given theory of counterfactuals against this. For example, Lewis (1979, 472) cites the example to motivate that in his condition (4), approximate similarities in particular matters of fact should count for a small amount, rather than nothing. Other cases, for example that given by Tichy (1976), seem to pull the other way; there is a literature on how to distinguish the cases, which I have no space to discuss here (see eg Shaffer (2004)).

In introducing the example, Slote considers the alternative line of regarding (M) as false. This is explicitly defended by Phillips (2007). It is tempting to think, argues Phillips, that since all the causally relevant factors are the same in the counterfactual situation in which the bet is made, the outcome will be the same. But to argue like this is just to assume determinism, *contra* the stipulation of indeterminism in the example (47).

What I want to suggest here is that there is no way to settle this dispute. On the one hand, we are inclined to evaluate the counterfactual by noting that the bet does not affect

the coin toss, and hence keep the toss's outcome fixed; this supports the truth of (M). On the other hand, we are also inclined to evaluate counterfactuals by fixing the history of the world up to the moment where the truth of the antecedent forces a divergence, and then imagine how things develop; since the coin toss is indeterministic, (M) is not true. The intuitions pull different ways, and cannot both be accommodated. Neither way seems wrong, casting doubt on the existence of a robust "default" method of evaluating counterfactuals, of the sort described by Lewis.

Similar remarks apply to the issue of *conjunction conditionalization*: does the truth of A and C entail the truth of A > C? On the one hand, it seems that it must. What better evidence could one have for A > C than that A actually turns out to be the case, and C as well? One might think this is the one case where one is really in a position to verify a subjunctive! Both the Lewis and the Stalnaker semantics validate this, since for both of them the actual world is the unique closest A-world, should A be true in the actual world.

On the other hand, suppose I am thinking of buying a ticket in a (large, indeterministic) lottery. I optimistically announce "if I were to buy a ticket, it would win". There is considerable pull to the thought that the subjunctive is simply false, even in the event that I do in fact buy a ticket and it does, in fact, win.

We do not have clear intuitions about even as simple a matter as whether A & C entails A > C. Nor is there any agreement on this amongst philosophers; see eg Walters and Williams (2013) for references. I am inclined once again to conclude that the issue is indeterminate in a strong sense: there is no way to settle the dispute.

5 Truth conditions and the Gibbard phenomenon

Alan Gibbard (1981, 231-2) gave a famous example (*Sly Pete*) which poses a problem for theories of *indicative* conditionals.[11] Since the example has some unnecessary complications, I give a simpler example which maintains the essential structure.

Suppose a crime was committed (by exactly one person) and that there are three suspects: the butler, the gardener, and the cook. At the time of the crime, Holmes could see the gardener, gardening; he is therefore able to assert

(H)   If it wasn't the butler, it was the cook.

But at the same time, Watson saw the cook, cooking, and so can assert

(W)   If it wasn't the butler, it was the gardener.

Of course, it was the butler whodunit. The puzzle is that (H) and (W) seem contradictory, yet neither Holmes nor Watson seems to be in any way mistaken. This seems to leave three ways out: (i) the conditionals are material (in which case they are both true); (ii) the conditionals have truth conditions, but these must be radically relativized to the speaker's epistemic states; and (iii) the conditionals do not have truth conditions. Gibbard produced his example as a way for arguing for (iii). He holds that indicative conditionals do not have truth conditions, whilst maintaining that counterfactual conditionals do.

Edgington (1991, 206-7; 1995, 295; 1997) seems to have been the first to notice that the "Gibbard phenomenon" can apply to counterfactuals as well; in fact, that "for any contingent conditional, the world may be such that the Gibbard phenomenon can arise".[12] If, after the case is closed, Watson asks Holmes "Why did you suspect the cook?", Holmes can felicitously reply: "Well, if it hadn't been the butler, it would have been the cook". Edgington draws the conclusion that there is no "ideal thing to think"; hence no objectivity and no truth value for in these cases. (She thinks no conditionals with false antecedents

---

11   There are earlier examples in the literature which turn on the same feature, going back as far as Lewis Carroll; see Rieger (2013, §5).

12   Van Fraassen (1980, 116-7) uses a Gibbard-like example (in fact, the one from Lewis Carroll) to argue for the essentially context-relative nature of conditionals; in reply, Stalnaker (1984, 151) complains that van Fraassen's examples are indicative, not subjunctive. But the argument carries over to the counterfactual case as well. See also Swanson (2013) for discussion of Gibbard cases involving subjunctives.

have truth values; in non-Gibbard cases there can be an objective "right thing to think", if there is an objective conditional probability. But this doesn't exist, she thinks, if the antecedent is ruled out by available facts. Her wider case that subjunctives have no truth values, which I shall not consider here, is made in most detail in her 2008.)

One might think that Edgington is too hasty in inferring from the facts ruling out the antecedent in the actual world to the conclusion that there is no objectivity in the counterfactual. Suppose, for example, that the cook is a gentle soul who would never harm anyone, but the gardener shared the butler's murderous intent, and indeed drew straws with him as to who should carry out the deed. This seems to favour "if it hadn't been the butler, it would have been the gardener" over the contrary counterfactual.[13]

On the other hand, suppose instead that both the cook and the gardener are incapable of harming anyone, but the butler jointly planned the murder with the chauffeur (a brute, who was away on the fatal night, but was determined the victim should die). Holmes might then also assert with felicity "if it hadn't been the butler, it would have been the chauffeur".

What seems right here is that there are sometimes two ways to hear a counterfactual, one an "epistemic" reading (in the sense of Gibbard 1981), in which it behaves much like an indicative, and one of the more usual "metaphysical" or "closeness" type (see for example Khoo (2015, §2.1)). Thus Holmes's conditional "if it hadn't been the butler, if would have been the cook" is correct on an epistemic reading, but not on a metaphysical one.

In any case, all this does not seem welcome news for those who favour scientific uses of counterfactuals, as it threatens to introduce a further axis of variation into the picture: for every conditional, we need to consider whether it might have be an epistemic reading that is intended. Maybe there is a way to winnow out the epistemic uses, though it is not obvious how this is to be done. And if Edgington is right, subjunctives are not truth-apt at all. She plays down the significance of this, since there can still be criteria of correctness – it is just that these are not truth and falsity. But a considerable amount of work remains to be done, for example in finding surrogates for entailment, and in explaining how conditionals behave in compounds.

---

13   This is essentially the point made by Morton (1997) about a different example of Edgington's.

6 Conditional Excluded Middle

Another example of a fundamental, but disputed, principle is the following. Do subjunctive conditionals obey the law of *Conditional Excluded Middle*, that is

(CEM)   (A > C) v (A > ¬C)?

Exploring one's intuitions with examples is inconclusive. On the one hand, a sentence like "either if it had rained the match would have been cancelled, or if it had rained the match would not have been cancelled" looks true. Perhaps, though, we hear it as "if it had rained, either match would have been cancelled or it would not", and hence trivially true. Natural language scope distinctions are notoriously confused in conditional contexts, as the case of negation shows.

On the other hand, symmetric examples pull the other way. Recall Quine's example above: if Bizet and Verdi had been compatriots, would they both have been French? Or both Italian? One hesitates to takes seriously the idea that exactly one of the counterfactuals is true, as appears to be required by CEM. Perhaps Quine's example leaves some small room for asymmetry,[14] but this is easily remedied: for example, we can consider the subjunctives "if I had tossed the coin, it would have come up heads" v "tails", said of an indeterministic coin.

On Lewis's semantics, CEM fails; for clearly if in some of the nearest A worlds, C is true, and some, C is false, neither A > C nor A > ¬ C is true. Stalnaker's position is more complicated. Since the selection function picks out a single possible world, in which C will be either true or false, it seems that CEM is validated. On the other hand, as noted above, Stalnaker admits that there is widespread indeterminacy in the selection function. So for example he regards the truth values of both the Verdi/Bizet counterfactuals as indeterminate. Does this mean that, after all, CEM fails? It does not, because Stalnaker favours a supervaluational approach (1980). On all ways of making determinate the indeterminate selection function, CEM is validated, and hence it comes out true in the semantics.

One might be concerned that Stalnaker has kept CEM in letter but not in spirit. After all, we have a situation where a disjunction is true while neither disjunct is. To put the point another way, CEM is preserved, but not bivalence. The worries here parallel those

---

14   Actually, Verdi's birth (in 1813, near Parma) took place in the First French Empire, so it might be reasonable to argue that a situation in which they are both French is much nearer than one in which they are both Italian.

concerning supervaluations in their more familiar setting of ordinary vagueness; excluded middle is "supertrue", but is supertruth really truth? [15]

There is a literature on CEM which I will not attempt to go into here. But the dispute provides another example of a fundamental issue about subjunctives about which there is widespread disagreement.

The situation seems *worse*, actually, than most examples of conceptual analysis. For example, in epistemology, there is no agreement about how to analyse the notion of knowledge. There does seem to be, though, a reasonable consensus about the intuitive verdict in each case. Similarly in ethics, with the analysis of right action. In these cases, the difficulty is fitting a theory which captures a reasonably robust collection of intuitions. With counterfactuals, the intuitions are themselves highly disputed.

I conclude, therefore, that the type of intersubjective agreement which Lewis and Stalnaker's approach seems to require seems unlikely to be forthcoming.

Perhaps, though, Lewis and Stalnaker's realism is too moderate. What if we reject all indeterminacy, and adopt a more hard-core realism?

7 The ultra-realists

Some, indeed have taken that route. Hawthorne (2005) discusses the following problem for possible worlds accounts. Consider a counterfactual such as "if I had dropped the plate it would have fallen to the floor". In the vast majority of worlds in which I drop the plate, it does indeed break. But, because of quantum theory, there is a small but non-zero chance that the plate will instead fly off sideways. And the worlds where this happens seem to be as close as the more common ones where the plate falls to the floor. Hence the counterfactual comes out false on Lewis's theory (and indeterminate on Stalknaker's), despite looking intuitively true.[16] Following a gloomy examination of the prospects of patching Lewis's theory to get the desired result, Hawthorne continues:

> In closing I might mention that my own preference is to opt for a picture according to which, for any possibility that P, and for any world w, there is a unique closest world to w where P.

---

15   See Williamson (1994, Ch. 5).

16   See Edgington (1995, 258), Bennett (2003, §96) and Hájek (2016).

A related proposal is to be found in Schulz (2014). Wishing to have the effect of having a single world picked out, but finding it implausible (1025) that there exists (in the style of Hawthorne or in the Stalnaker semantics) a *unique* world relevant to the evaluation of a counterfactual, Schulz gives a modification of Lewis's semantics which replaces the universal quantifier with Hilbert's epsilon-symbol. This has the effect of picking out an *arbitrary* world from amongst those that are candidates. The result, as with Hawthorne, is that the "dropped plate" counterfactual should be given a credence which is very high but falls short of 1, reflecting the fact that there is a minute probability that the chosen world will be one in which the plate flies off to the side.

Schulz leaves open the metaphysics of the selection function, but leans towards endorsing a realist stance according to which "there simply is a distinguished but arbitrary selection function which supplies epsilon-terms with referents".

Both Hawthorne's and Schulz's approaches naturally endorse CEM, since either A > C or A > ¬C will be true, depending on whether C or ¬C is true in the unique selected world. A paper by Jeremy Goodman (2015b) puts CEM centre-stage. Goodman argues that CEM (which he takes to be true) entails various surprising metaphysical consequences, including the falsity of physicalism. To give an idea (slightly simplified) of the argument: Goodman considers a Max Black-style universe which consists only of two iron spheres rotating around each other. From CEM, it follows, he argues, that only one of the spheres has the property of being the one that would have been heavier, had they had different masses. And this property is not a physical one, thus showing that physicalism is false. Other heavyweight metaphysical theses which Goodman also derives are the identity of indiscernibles and the failure of various supervenience theses.

All three authors thus have a robustly realistic attitude to counterfactuals, on which there is no indeterminacy and counterfactuals are "brutely" true. How should we assess this? One is tempted merely to offer an incredulous stare, but can one do better?

Goodman's argument is surely asking for a *modus tollens*. If these drastic theses follow from taking such a robustly realistic view, it seems more plausible to doubt the premises than swallow the conclusion.

Nor is the absence of indeterminacy in accord with common sense. I take it that the reason the joke in the epigraph is funny is because the folk theory of counterfactuals makes it absurd that there should be a fact of the matter about the food preferences of a merely possible sibling. Indeterminacy is deeply ingrained in our pre-theoretical views.

Speaking more theoretically, this extreme realism runs into the same kind of problems as epistemicism in the case of vagueness (of which it is at least a cousin – perhaps more, if one views the source of indeterminacy of counterfactuals as arising entirely from vagueness). It is very hard to see *what could make it the case* that there is a determinate individual in some particular world whose food preferences determine the truth value of the conditional. In the case of epistemicism about ordinary vagueness, there is at least the outline of an answer of what determines, say, the border between blue and green: that ultimately it depends on the usage and verdicts of speakers, though in some highly complex way that remains impenetrable to us. But it does not seem credible that our linguistic practices fix a determinate truth value for every subjunctive conditional.

8 Conclusion

Are counterfactuals suitable for use in scientific purposes? We can agree with Quine that they may have a heuristic purpose in introducing dispositional terms. But can they play some more substantial role? If we can put aside the extreme realist position as incredible, the question is whether some more moderate realist position can be made to work. Such a theory admits indeterminacy in general, but holds that it can be resolved in appropriate cases, or at least a high proportion of them.

I do not claim to have established that this cannot be done. But I hope to have shown that even the modest realism of Lewis and Stalnaker faces considerable obstacles. In place of the approximate consensus which would give some support to their position, we find little agreement about subjunctives or their logic. And the voluminous recent work on conditionals has made things worse rather than better.

Clearly counterfactuals should not be eschewed completely. As well as the Quinean heuristic use and their obvious use in informal reasoning, they will undoubtedly continue to have philosophical applications. Analysing some notion in terms of counterfactuals may bring clarity. It may also give the appearance of objectivity, that some mind-independent truths are being uncovered. But the objectivity is spurious. That counterfactuals are essential in analysing a concept is, rather, a sign that it is intelligible only relative to human concerns.[17] At best we can hope for intersubjective agreement; at worst even that will not be forthcoming.

---

17   For example, van Fraassen (1980, 118) draws this conclusion about the concept of *explanation*: it makes essential use of counterfactual language, and is therefore

It seems, then, that Quine's strictures on counterfactuals were broadly correct.   The right course of action with the subjunctive conditional is to "set it outside the systematic fabric of science".[18][19]

Adam Rieger

University of Glasgow

References

Bennett, J., 2003. *A philosophical guide to conditionals*. Oxford University Press.

Edgington, D., 1991. The mystery of the missing matter of fact. *Aristotelian Society Supplementary Volume*, *65*, 185-209.

Edgington, D., 1995. On conditionals. *Mind*, *104*, 235-329.

Edgington, D., 1997. Truth, objectivity, counterfactuals and Gibbard. *Mind*, *106*(421), 107-116.

Edgington, D., 2008. Counterfactuals. *Proceedings of the Aristotelian Society* (108), 1-21.

Fine, K., 1975. Review of Lewis (1973). *Mind*, *84,* 451-8.

Van Fraassen, B., 1980. *The Scientific Image.* OUP.

---

context-dependent rather than scientific. *Causation* is another likely candidate, as previously mentioned in connection with Lewis.

18   Quine (1995, 98).

19   I was prompted to think about conditionals again when I attended the very enjoyable 2nd Belgrade Workshop on Conditionals in May 2016; in particular this paper has been influenced by talks I heard there given by Lee Walters, on the Gibbard cases, and Alan Hájek, on "counterfacts". Back in Glasgow, Alan Weir read a draft and gave me detailed and invaluable comments about Quine and counterfactuals. It was a particular pleasure to be able to discuss philosophy with him again. Gary Kemp also read a previous version and made extremely helpful comments.

Gibbard, A., 1981. Two recent theories of conditionals. In Harper et all (1981, 211-247). Springer Netherlands.

Goodman, J., 2015a. Counterfactuals and comparative similarity. MS at http://users.ox.ac.uk/~newc3660/CounterfactualsWithoutCloseness.pdf, accessed 23 Nov 2016.

Goodman, J., 2015b. Consequences of conditional excluded middle. MS at http://users.ox.ac.uk/~newc3660/ConsequencesCEM.pdf, accessed 30 Nov 2016.

Goodman, N., 1970. Seven strictures on similarity, in Foster, L. and Swanson, J. W., *Experience and Theory,* University of Massachusetts Press.

Hájek, A., 2016. *Most counterfactuals are false.* MS at http://philrsss.anu.edu.au/people-defaults/alanh/papers/MCF.pdf, accessed 30 Nov 2016.

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

Hawthorne, J., 2005. Chance and counterfactuals. *Philosophy and Phenomenological Research*, *70*(2), 396-405.

Khoo, J., 2015. On indicative and subjunctive conditionals. *Philosophers' Imprint* 15(32).

Lewis, D., 1973. *Counterfactuals*. John Wiley & Sons.

Lewis, D., 1974. Causation. *The Journal of Philosophy (70), 556-567.*

Lewis, D., 1979. Counterfactual dependence and time's arrow. *Noûs*, 455-476.

Morton, A., 1997. Can Edgington Gibbard counterfactuals? *Mind*, *106*(421), 101-105.

Phillips, I., 2007. Morgenbesser cases and closet determinism. *Analysis 67* (293), 42–49.

Quine W., 1960. *Word and Object*. MIT Press.

Quine, W., 1974. *The Roots of Reference.* Open Court.

Quine, W., 1982. *Methods of Logic*. Harvard University Press.

Quine, W., 1995. *From Stimulus to Science*. Harvard University Press.

Rieger, A., 2006. A simple theory of conditionals. *Analysis*, *66*(291), 233-240.

Rieger, A., 2013. Conditionals are material: the positive arguments. *Synthese*, 190(15), 3161-3174.

Rieger, A., 2015. Defending a simple theory of conditionals. *American Philosophical Quarterly,* 52(3), 253–60.

Schaffer, J., 2004. Counterfactuals, causal independence and conceptual circularity. *Analysis*, 64(284), 299-308.

Schulz, M., 2014. Counterfactuals and arbitrariness. *Mind*, *123*(492), 1021-1055.

Slote, M., 1978. Time in counterfactuals. *The Philosophical Review*, *87*(1), 3-27.

Stalnaker, R., 1968. A theory of conditionals. In Harper et al, (1981, 41-55).

Stalnaker, R., 1980. A defense of conditional excluded middle. In Harper et al, (1981, 41-55).

Stalnaker, R., 1984. *Inquiry*. MIT Press.

Swanson, E., 2013. Subjunctive biscuit and stand-off conditionals. *Philosophical Studies*, 163(3), 637-648.

Tichý, P., 1976. A counterexample to the Stalnaker-Lewis analysis of counterfactuals. *Philosophical Studies*, *29*(4), 271-273.

Walters, L. and Williams, J.R.G., 2013. An argument for conjunction conditionalization. *The Review of Symbolic Logic*, *6*(04), 573-588.

Williamson, T., 1994. *Vagueness.* Routledge.