Erkenn (2016) 81:705–719
DOI 10.1007/s10670-015-9763-9

ORIGINAL ARTICLE

Conjunction, Connection and Counterfactuals

Chaoan He[1]

Received: 5 June 2014 / Accepted: 22 August 2015 / Published online: 3 September 2015

Abstract The standard Lewis–Stalnaker semantics of counterfactuals, given the Strong Centering Thesis, implies that all true–true counterfactuals are trivially true. McGlynn (Analysis 72:276–285, 2012) developed a theory, based on Penczek (Erkenntnis 46:79–85, 1997), to rehabilitate the non-triviality of true–true counterfactuals. I show here that counterfactuals with true but irrelevant components are counterexamples to McGlynn's account. I argue that an extended version of the connection hypothesis is sustainable, and grounds a full theory of counterfactuals explicable in a broadly standard way, if an indispensable asymmetry between semifacuals and other counterfactuals is acknowledged.

1 Introduction

The standard Lewis–Stalnaker semantics defines a counterfactual $\Phi \Box\!\!\rightarrow \Psi$ as being true if and only if one of the following two conditions is satisfied:

(i) there is no Φ-world accessible to *i*,
(ii) some (Φ & Ψ)-world is closer to *i* than any (Φ &~Ψ)-world.

If the so-called Strong Centering Thesis is assumed, however, the standard semantics implies some intuitively undesirable results.

✉ Chaoan He
chaoanhe@hotmail.com

[1] School of Humanities, Donghua University, No. 2999, North Renmin Road, Songjiang District, Shanghai 201620, People's Republic of China

Strong Centering: A world is closer to itself than any other world is to it.

When Φ and Ψ are true at world *i*, clause (ii) is trivially satisfied. For *i* is a (Φ & Ψ)-world, and it is closer to itself than any other world whatsoever, let alone any (Φ &∼Ψ)-world. So the conjunction of the standard semantics and Strong Centering validates what has been called Conjunction Conditionalization:

CC: $(\Phi \,\&\, \Psi) \supset (\Phi \mathrel{\Box\!\!\rightarrow} \Psi)$

That seems to be a defect of the standard semantics, for it leads to the unexpected consequence that all true–true counterfactuals (counterfactuals with true antecedents and true consequents) are trivially true. As far as I can see, there are two typical types of true–true counterfactuals which constitute putative counterexamples to CC. First, there are cases where the antecedent is relevant to, but insufficient for, the consequent, even under certain background conditions. McDermott's coin case (McDermott 2007) serves to illustrate. A coin was to be tossed twice. I bet on "two heads" and two heads indeed came up. The standard semantics implies that the following true–true counterfactual is true:

(1) If at least one head had come up, I would have won.

But (1) appears to be intuitively false, for if only one head had come up, I would have lost. Take another example. Suppose that Oswald is the actual killer of Kennedy, then (2) would be a true–true counterfactual.

(2) If someone killed Kennedy, then Oswald would have been guilty of killing Kennedy.

There is some pull to count (2) as false, though the standard semantics says that it is trivially true.

Second, there are true–true counterfactuals where the antecedent is completely irrelevant to the consequent.

(3) If Casper had come to the party, then it would have rained.

Since Casper's attending the party, one may wish to say, *has nothing to do with* the raining, (3) is intuitively false. But the standard semantics once again classifies it as a trivial truth.

There have been many attempts to remedy this defect of the standard semantics. Though CC itself has received a couple of impressive defenses from some able philosophers (Walters 2009, 2015; Walters and Williams 2013), many believe (Penczek 1997; Gundersen 2004; McGlynn 2012; Cogburn and Roland 2013) that inessential revisions could be incorporated into the standard semantics to invalidate CC and to accommodate the troublesome cases.

The straightforward move of giving up Strong Centering and replacing it with Weak Centering, it should be noted, has been rightly discarded on solid grounds.

Weak Centering: A world is no less close to itself than any other world is to it.

As Walters (2015) and McDermott (2007) rightly observe, adopting Weak Centering is not fully satisfactory in the end. For there are equally unproblematic cases of counterfactuals that seem to require Strong Centering. As a consequence, a variety of revisionary semantics have been proposed that retain Strong Centering.

In this article, I argue that one such revisionary semantics recently developed by McGlynn (2012), based on Penczek (1997), is not successful in its full generality. And I propose a novel account of counterfactuals that rehabilitates Penczek's essential idea of the connection test, grounded upon the following thesis which one may call the connection hypothesis:

> CH: Counterfactuals suggest the presence of (usually causal or explanatory) connections between their antecedents and consequents, and are true if and only if the connections obtain.

McGlynn's account, it is worth pointing out at the outset, has already received criticisms on a number of grounds. Walters, for instance, argues that McGlynn's account is "radically incomplete at best" (forthcoming, p. 18), and it "invalidates a range of compelling principles of counterfactual logic" (forthcoming, p. 19).

This article takes a different route. I attempt to show that true–true counterfactuals with irrelevant components are counterexamples to McGlynn's account. McGlynn (2012, p. 282) believes that his semantics is such a minor revision of Penczek's that the strength of the latter should "carry over virtually unchanged" to his own. I argue that the revision is a rather major one. McGlynn's remedy actually abandons the one essential idea of the connection test underlying Penczek's proposal. And this major revision leaves McGlynn's account vulnerable to counterexamples in ways that Penczek's is not.

I'll start, in Sect. 2, by introducing what I take to be counterexamples to McGlynn's semantics. I also consider in brief along what lines a new semantics for counterfactuals could be developed. I argue, in Sect. 3, that counterfactuals such as (3) are indeed false and constitute genuine counterexamples to McGlynn's account. In the process, I develop an extended version of the connection hypothesis, and I show that this hypothesis is not only sustainable on its own, but also capable of rehabilitating Penczek's connection-test and grounding a full theory of counterfactuals.

2 Conjunction Conditionalization and McGlynn's Semantics

One earlier attempt to invalidate CC without giving up the fundamental ideas of the standard semantics is from Penczek (1997). He suggests that the truth conditions of counterfactuals are captured by the following definition:

$\Phi \,\Box\!\!\rightarrow \Psi$ is true at world i if and only if EITHER

(i) there is no Φ-world accessible to *i*; OR:
(ii) some (Φ & Ψ)-world is closer to *i* than any ($\Phi \& {\sim}\Psi$)-world, and
(iii) $({\sim}\Phi \& {\sim}\Psi) > (\Phi > \Psi)$.

Here $\Phi > \Psi$ is the Lewis Conditional interpreted in the standard Lewis way:

LC: $\Phi > \Psi$ is true at world *i* if and only if either

(i) there is no Φ-world accessible to *i*, or
(ii) some (Φ & Ψ)-world is closer to *i* than any ($\Phi \& {\sim}\Psi$)-world.

The idea behind the introduction of clause (iii), as Penczek (1997, p. 81) sees it, is that it should not count in a counterfactual's favor if its antecedent and consequent merely happen to be true, and we have to force ourselves to consider what kind of connection might obtain between its antecedent and consequent when both are *false*.

It is obvious that PC delivers intuitively correct verdicts for true–true counterfactuals such as (1). Consider the closest (${\sim}\Phi \& {\sim}\Psi$)-worlds, where no head came up and I lost. Clause (iii) effectively takes us to consider whether there are any worlds where some heads came up and I won which are closer to the (${\sim}\Phi \& {\sim}\Psi$)-worlds than any worlds where some heads came up but I lost. Obviously those worlds where only one head came up and I lost are no less close, if not closer, than those where two heads came up and I won.

Penczek's account, though appealing in cases such as (1), runs into problems with semifactuals, i.e., counterfactuals with false antecedents and true consequents. Suppose that John enjoys a fame of being an annoying fellow. He was late for an out-door party, and it rained shortly after his presence. Jill, being disappointed with this, said to her friend: "John shouldn't have been here, he brought the bad weather which spoils the party". Her friend, having no prejudice against John, responded:

(4) If John hadn't been here, it would still have rained.

The (${\sim}\Phi \& {\sim}\Psi$)-worlds in this case are such that John was at the party and it didn't rain. But there surely are worlds where John was not there and it didn't rain which are no less close to the (${\sim}\Phi \& {\sim}\Psi$)-worlds than worlds where John was not there and it rained. Hence (4) comes out false according to PC, contrary to what it appears to be.

Why Penczek's account gets this wrong is not hard to diagnose. His clause (iii) invites us to consider what kind of connection might obtain between the components of (4), but nothing really guarantees that (4) has to be true in virtue of the existence of *any* connection at all. On the contrary, (4) seems to be true rather in virtue of the *absence* of a certain connection, namely, John's not attending the party would have undermined the rain weather. This defect of his account, as McGlynn (2012, p. 280) rightly observed, originates from Penczek's supposition that counterfactuals suggest connections in virtue of their form. Though appealing at a first glance, this supposition is simply too bold to be true.

Given the truth of semifactuals like (4), McGlynn's own suggestion instead is that we forget about the idea of the connection test, and substitute Penczek's clause (iii) with

(iii′) $(\sim\Phi) > (\Phi > \Psi)$.

This new proposal does escape the trouble posed by semifactuals. As clause (iii′) requires the consideration of the truth of $\Phi > \Psi$ under the supposition that Φ is false, it becomes a matter of determining the truth value of $\Phi > \Psi$ relative to the actual world, when Φ is *actually* false. McGlynn's clause (iii′), one should note, simply collapses to that of Lewis on conditionals with false antecedents. As (4) is a semifactual, it is to be treated in the standard Lewis way. That is, one only needs to check if there are worlds where John didn't attend the party and it didn't rain that are at least as close to the actual world as worlds where John didn't attend the party and it rained. Obviously there are no such worlds. This is because, given the actual rain weather, it would take a *greater* departure from actuality to move to a non-rain world. And that means that (4) is true. McGlynn's account, therefore, is doing a better job than Penczek's on semifactuals.

Not only so. McGlynn's account *seems* to be doing an equally good, if not better, job on the troublesome true–true counterfactuals which constitute putative counterexamples to CC. Consider (1):

(1) If at least one head had come up, I would have won.

Since the $\sim\Phi$-worlds are such that no head came up, the closest worlds to them are ones where one head came up.[1] But in these closest worlds, I would have lost. Thus, (1) is correctly rendered as false under McGlynn's account.

McGlynn believes that his account, differing from Penczek's in such a minor fashion, is exactly what it takes to avoid PC's vices while retaining its virtues. The following analysis, however, should suffice to show that it is not. As mentioned above, there seems to be two typical types of true–true counterfactuals which constitute putative counterexamples to CC. First, there are cases where the antecedent is relevant to, but insufficient for, the consequent; second, there are cases where the antecedent is completely irrelevant to the consequent. It is for this second type of cases, it will become clearer, that McGlynn's account fails to provide correct verdicts, or so I shall argue.

Let's consider the true–true counterfactual

(3) If Casper had come to the party, then it would have rained.

[1] Some philosophers might find this disputable. Pollock, for instance, thinks that "the idea of the magnitude of change is not the same as that of the comparative similarity of worlds" (Pollock 1976, p. 21). He would probably argue that a change yielding two heads up is no greater dissimilarity-making than one yielding only one head up. Nothing of substance here, however, hinges on that. For whatever the ordering of similarity one may favor, worlds where one head came up and I lost are certainly *no less* close to the actual world than worlds where two heads came up and I won.

According to McGlynn's account, we have to consider the $\sim\Phi$-worlds, where Casper didn't attend the party. We are then to determine the truth value of (3′) in these worlds:

(3′) Casper comes to the party > it rains.

To do so, we go to the closest worlds v to the $\sim\Phi$-worlds, in which Casper came to the party. Now, are there v worlds in which it didn't rain that are at least as close to the $\sim\Phi$-worlds as the v worlds where it rained? Due to the reason similar to that regarding (4), it is safe to say that there are no such v worlds. For it would take a *greater* departure from actuality to move from a rain world to a non-rain world than from a rain world to a rain world. And that means, (3) is true according to McGlynn's account. Since Casper's coming to the party, however, *has nothing to do with* the raining, (3) is intuitively false. If, as I shall argue in greater detail, true–true counterfactuals with irrelevant components are indeed false, cases such as (3) constitute genuine counterexamples to McGlynn's semantics.

But before going into that, let me say something about the general lines along which a more promising semantics for counterfactuals may be developed, drawing the morals here emerged from the present discussion. As one may easily check for herself, a similar difficulty with true–true counterfactuals with irrelevant components does not arise for Penczek. Penczek's problem, therefore, seems to remain only with semifactuals. The aim of McGlynn's semantics is exactly to avoid this particular problem, while preserving all the advantages of Penczek's account. So to achieve what McGlynn envisages, we'd better go back to Penczek's problem with semifactuals.

Regarding the problem with semifactuals, Penczek (1997, p. 84) once recommended that we could either explore further modifications of the traditional semantics or treat semifactuals as a distinct type of conditional deserving separate semantics of their own. The unexpected troubles for Penczek' account with semifactuals, despite its success in nearly all other cases, is itself a fairly strong indication that semifactuals may be distinct from other species of counterfactuals in some sense. Accommodating semifactuals without giving them their due distinction, as Penczek and McGlynn do, may generate unexpected problems. Yet, as we shall see in the next section, this distinction is not so substantive as to make semifactuals a *sui generis* kind of conditionals other than counterfactuals. I suggest that the distinction is better seen as an intra-kind one within the family of counterfactuals rather than an inter-kind one between counterfactual conditionals and another kind of conditionals. This is because, all subspecies of counterfactuals, semifactuals included, can be said to rest on a single key notion, the notion of connection.

I propose that a distinction be made between *connection-affirming* counterfactuals and *connection-denying* ones. At the bottom, semifactuals are connection-denying, and all other counterfactuals are connection-affirming. As Gundersen (2004, p. 7) writes on semifactuals, "reasoning counterfactually, we are sometimes investigating whether some 'counterfact' would bring such a 'disturbance' into the actual course of events so as to prevent some fact from occurring", the connection denied in a semifactual is that between a counterfact and the prevention of some fact.

So it is tempting to define the truth condition of a counterfactual as a disjunction of two possibilities, giving the distinction its due.

$\Phi \square\!\!\rightarrow \Psi$ is true at world i if and only if EITHER

(i) world i is a Φ-world or a $\sim\Psi$-world, and $\Phi \square\!> \Psi$, OR
(ii) world i is a $(\sim\Phi \,\&\, \Psi)$-world, and $\sim(\Phi \square\!> \sim \Psi)$.

Here $\Phi \square\!> \Psi$, call it the Penczek Conditional, is defined in the Penczek way:

PC: $\Phi \square\!> \Psi$ is true at world i if and only if EITHER

(i) there is no Φ-world accessible to i; OR:
(ii) some $(\Phi \,\&\, \Psi)$-world is closer to i than any $(\Phi \,\&\!\sim\!\Psi)$-world, and
(iii) $(\sim\Phi \,\&\!\sim\!\Psi) > (\Phi > \Psi)$.

In effect, our proposed clause (i) defines the truth condition of connection-affirming counterfactuals, and clause (ii) defines that of connection-denying counterfactuals. Whereas how the distinction between connection-affirming/connection-denying should be drawn among counterfactuals is a question that will be addressed in detail in the next section.

Penczek's account works for all subspecies of counterfactuals other than semifacuals, and McGlynn's works for all subspecies of counterfactuals other than true–true counterfactuals with irrelevant components. If viable, the semantics here proposed has the potential of preserving the virtues of the accounts of both Penczek and McGlynn, while avoiding their fatal flaws. On the one hand, this new account, unlike that of Penczek, does not test the *presence* of connections between the components of semifactuals, rather it tests the *absence* of certain reverse connections. Take our previous example to illustrate,

(4) If John hadn't been here, it would still have rained.

Since (4) is a semifactual, it is taken care of by our clause (ii). The description of the relevant scenario implies that, literally, what (4) says is that John's not attending the party would not have prevented it from raining, or

(4′) NOT: If John hadn't been here, it would not have rained.

As our clause (ii) is further defined in terms of the Penczek Conditional, forgetting about the outermost negation operator "NOT" for the moment, we need first consider the $(\sim\Phi \,\&\!\sim\!\Psi)$-worlds in which John attended the party and it rained. Then we go to the closest worlds v to them in which John didn't attend. Obviously, the v worlds are all such that it rained there. This means that "If John hadn't been here, it would not have rained" is false, hence its negation, i.e., (4′) and (4), are true, as they intuitively seem to be.

On the other hand, the semantics here proposed, unlike that of McGlynn, does not abandon Penczek's core idea of the connection test, when it comes to the troublesome true–true counterfactuals. The total strength of Penczek's account, it is safe to say, should be carried over unchanged for true–true counterfactuals with irrelevant components.

So there seems to be, from my perspective, three desiderata for any accurate account of counterfactuals. First, the falsity of true–true counterfactuals with insufficient antecedents such as (1), second the falsity of true–true counterfactuals with irrelevant components such as (3), and third the truth of semifactuals with irrelevant components such as (4). Whereas Penczek's account secures the first two, McGlynn's secures the first and the third, my own account captures all three. In the next section, I proceed to show why true–true counterfactuals with irrelevant components constitute genuine counterexamples to McGlynn's account, and how my preferred semantics could be justified on independent grounds.

3 The Connection Hypothesis and Counterfactuals with True Components

There are readers, I suspect, who would be unmoved by the putative counterexamples we put forward against McGlynn'a account *per se*. Indeed, they may be puzzled. Why, one may wonder, should anyone believe that (3) is false at all? Why not accept the truth of (3) *tout court* in the first place? This skepticism is certainly justified, and I shall address it here.

So why does anyone believe that (3) is false? The main reason, as indicated in the previous discussions, is reflected in the connection hypothesis:

> CH: Counterfactuals suggest the presence of (usually causal or explanatory) connections between their antecedents and consequents, and are true if and only if the connections obtain.

In the case of (3), the antecedent and consequent are not connected, or not in an appropriate way. John's attending the party neither causes nor explains the rain. The rain has nothing to do with Casper's attendance. It is simply irrelevant. And it is the absence of the alleged connection that grounds the very intuition that (3) is false, an intuition we favor. So if CH or some similar principle is plausible, there is a good reason to believe that (3) is false.

Appealing to a layman as it may be, this line of thought is far from convincing to a counterfactual theorist. As Pollock notes:

> There is a more or less traditional assumption about subjunctive conditionals that has been uniformly rejected in the recent literature. This is the assumption that a subjunctive conditional asserts the existence of a *connection* between the antecedent and consequent. Certainly some simple subjunctives are true because such a connection exists, but this is not invariably the case. The existence of such a connection is a sufficient, not a necessary, *condition* for the truth of a simple subjunctive. (Pollock 1976, p. 25)

The popular rejection of CH, as we find it in the literature, has its source in the observation that a counterfactual can well be true if no such a connection obtains. We might say of a witch doctor, Pollock's example goes, that "it would not rain if he did not do a rain dance, but it would not rain if he did either". (Pollock 1976, p. 26) It is the *lack* of a connection, rather than the presence of one, that the sentence seems to express. Thus not all counterfactuals are true in virtue of the presence of connections. If CH was indeed untenable, the straightforward move of counting (3) as false would have been significantly undermotivated. Though that is far from enough to establish in turn the truth of (3) in a decisive way, it certainly lends credence to the contention that (3) *might* be true after all.

I am of perfect agreement with our disputant's observation that CH is untenable *per se*. But I have a different opinion on what should be made out of that observation. Scarcely anyone denies that *some* counterfactuals do suggest the existence of certain connections, and are true in virtue of the existence of such connections. Laws of nature, to exploit one such instance, are commonly said to support counterfactuals. When a counterfactual is employed to reveal a causal or nomic connection endowed by laws of nature, its truth hinges exactly on the existence of the connection in question.

So the Pollock observation really suggests that CH might be telling only *half* a story. If we embrace the more flexible idea that counterfactuals express *the presence or the absence* of certain connections, a version of connection hypothesis may prove sustainable after all. To employ Pollock's example again: a witch doctor didn't do a rain dance, and (obviously not as a result) it didn't rain. The following counterfactual is certainly not to *affirm* that the witch doctor's rain dance would bring about not raining, rather it is to *deny* that her dance would undermine the non-rain situation:

(5) If the witch doctor did a rain dance, it would not have rained.

Since (5) is an average semifactual, and cases like it abound, it is very much desirable to build the thesis of connection-denial into the connection hypothesis, if the essence of counterfactuals is to be fully captured. In an examination of similar issues, Gundersen (2004, p. 8) once argued forcefully that there is an often ignored distinction between false–false counterfactuals and semifactuals, in that the former assert a "counterfactual dependence" between the antecedents and the consequents, and the latter deny a "counterfactual dependence" between the antecedents and the negation of the consequents. I find that idea very convincing and lovely, if the notion of counterfactual dependence is charitably understood as that of connection. I suggest, therefore, that the following extended version of the connection hypothesis would do a greater justice to a full understanding of counterfactuals:

> ECH: Either counterfactuals suggest the presence of (usually causal or explanatory) connections between the antecedents and consequents, and are true if and only if the connections obtain; or counterfactuals suggest the absence of such connections between the antecedents and the negation of the consequents, and are true if and only if the connections fail to obtain.

A question may immediately recommend itself concerning the bivalence here. What justifies the "either…or" classification of counterfactuals? Is there not "a third" variant of counterfactuals which neither affirm nor deny connections? I should confess that I have no ready arguments for the bivalence. But no one, as far as I know, has really suggested otherwise. Indeed, even the most obstinate opponents of CH seem to be committed to the bivalence. Pollock, for instance, writes:

> It seems that there are basically two ways that a simple subjunctive can be true. On the one hand, there can be a connection between the antecedent and consequent so that the truth of the antecedent would bring it about, i.e., necessitate, that the consequent would be true. On the other hand, a simple subjunctive can be true because the consequent is already true and there is no connection between the antecedent and the consequent such that the antecedent's being true would interfere with the consequent's being true. (Pollock 1976, p. 26).

What is more pressing a problem in this connection is how the connection-affirming/connection-denying distinction among various subspecies of counterfactuals should be drawn, particularly when it comes to the troublesome true–true counterfactuals. We shall return to this issue very soon. For the time being, let us say as a first approximation that false–false counterfactuals affirm connections, while semifactuals deny them.

On the one hand, the established practices of counterfactual theories regard false–false counterfactuals as *the* standard counterfactuals. The fact that these standard counterfactuals, in virtue of their form, seem to suggest the existence of certain connections, is presumably the principal motivation for the introduction of the naive connection hypothesis CH. Though CH, taken as a principle governing *all* counterfactuals across the board, is inadequate, its inadequacy is merely due to the existence of semifactuals and similar cases like Pollock's witch doctor. That ordinary false–false counterfactuals suggest the presence of connections is itself an innocent thesis worth preserving, as long as we do not overlook the fact that not all counterfactuals are of the false–false type.

On the other hand, semifactuals, as McGlynn (2012, pp. 280–281) himself put it, "are typically taken as denying that there's any such connection between their antecedent and their consequent". Virtually all denouncements of CH at market, in one way or another, turn on the indisputable observation that semifactuals deny, rather than affirm, certain connections. So the thesis that semifactuals are connection-denying is a consensus shared by the friends and enemies of CH alike.

Besides all this, there is a special characteristic of semifactuals which is a fairly strong indication that semifactuals express denials of certain connections. One could hardly fail to notice that semifactuals are naturally wedded to an expression involving the "even if…still" construction. To wit, the witch doctor case (5) is more naturally phrased as

(5′) *Even if* the witch doctor did a rain dance, it would *still* not have rained.

The "even if…still" construction seems to be naturally reserved for conditionals with false antecedents and true consequents. For the phrase "even if Φ", in a typical conversational context, more or less implies "actually not Φ", and "still Ψ", in a similar fashion, implies "actually Ψ". The word "even", as Lycan (2001) puts it, "carries a strong connotation having to do with contextual presumptions or expectations and circumstances' contravening those expectations". When saying that "even Grannie was sober", one indicates that "the level of sobriety of some gathering was so great that Grannie, whom one would not expect to be sober, was sober". And "not a creature was stirring, not even a mouse" suggests that "the stillness on the occasion in question was so great as to include all the mice—animals that one would normally expect to be moving about" (Lycan 2001, pp. 93–94).

As to "even if…still" conditionals like "even if Φ, still Ψ", similarly, there seems to be the contextual expectation that Φ would bring it about that not-Ψ, and the point of those conditionals is exactly to contravene such expectations. In the case of (5′), for instance, there is the contextual expectation that the witch doctor's rain dance would bring it about that it rains. And what (5′) seems to expresses is a contravention of that expectation, or, to put it in our terms, a denial of the expected connection. Such an expectation-contravention mechanism suggests that the connection denied in a semifactual, as here illustrated, is that between the antecedent and the negation of the consequent. Semifactuals "$\Phi \mathrel{\Box\!\!\rightarrow} \Psi$", therefore, can be squarely taken to be equivalent to "$\sim(\Phi \mathrel{\Box\!\!\rightarrow} \sim\Psi)$".

Now, what about the troublesome true–true counterfactuals? Should they be assimilated to semifactuals as connection-denying conditionals? The *actual* truth of Φ and Ψ, anyway, is automatically sufficient for denying the reverse connection between Φ and not-Ψ. If true–true counterfactuals were interpreted as connection-denying this way, one should notice, they would then be automatically true regardless of the specific contents of their components. McGlynn himself, indeed, as he indicates in a personal correspondence, is ready to meet our challenge by embracing the idea of employing the "even if…still" maneuver to rephrase true–true counterfactuals with irrelevant components and to explain away our sense that cases such as (3) are false. But an immediate result emerging from the previous consideration would suggest that true–true counterfactuals have to be classified as connection-affirming conditionals.

Consider, what would true–true counterfactuals look like under the expectation-contravention interpretation, if they were taken to be connection-denying. Suppose that (5) is a true–true counterfactual which denies a certain connection. According to the expectation-contravention interpretation, there should be the contextual expectation that the witch doctor's rain dance should bring it about that it rains. But that expectation simply makes no sense, given the background assumption that the witch doctor *did* do a rain dance and it didn't rain. There is just no room to expect that the witch doctor's rain dance would bring it about that it rains when he *actually* danced without bringing the rain. That is very different from the case in which (5) were a semifactual: when the witch doctor didn't do a rain dance, expecting his dance would bring it about that it rains makes perfect sense.

Since no similar expectations arise in (5) when it is taken as a true–true counterfactual, a contravention of some such expectation is not what (5) expresses either. Without the expectation-contravention mechanism working behind, it is difficult to see how true–true counterfactuals can be naturally rendered as connection-denying in the way semifacutals are. Whether there are other ways counterfactuals can be taken to suggest denials of connections is a question open to further disputes, one that I'm not in a position here to cast deeper. But given that semifactuals are the standard connection-denying conditionals, the failure on the part of true–true counterfactuals to work in this standard way lends considerable credence to the conclusion that true–true counterfactuals are not connection-denying at all.

This, however, might not be the end of the story. There could be a charge of an unwanted asymmetry. Since false–false counterfactuals are *the* standard counterfactuals, all other subspecies are non-standard or derivative. Semifactuals and true–true counterfactuals, being equally non-standard, should be treated on a par. But if semifactuals are classified as connection-denying while true–true counterfactuals as connection-affirming, there generates a drastic asymmetry.

If the antecedent of (5) is false, (5) is a semifactual which denies that the witch doctor's rain dance would undermine the non-rain weather; whereas if the antecedent is true, (5) is a true–true counterfactual which states that the witch doctor's rain dance would bring it about that it doesn't rain. Obviously, (5) is true in the former case and false in the latter. That almost amounts to saying that the truth value of (5) hinges exactly on that of its antecedent, given the truth of the consequent. But, as Walters (2015, p. 35) notes, "whilst people have thought that the truth of the antecedent should not count in favor of a counterfactual being true, they have not thought that it should count against it being true", which is effectively what we seem to be doing here.

In response to this incisive charge, it should suffice to point out that the truth value of the antecedent does not count in favor of or against *a counterfactual's being true*, rather, it only counts in favor of or against *classifying a counterfactual as being of one type than another*. Though there is a way of *tracing* the truth value of the counterfactual back to that of its antecedent, the truth value of conditionals such as (5) does not hinge on that of its antecedent *per se*. It only (partially) hinges on the specific type the counterfactual falls into, which the antecedent's truth value serves to determine. And a counterfactual's truth value itself is determined ultimately by the presence or absence of the alleged connection the counterfactual is taken to suggest.

There is a further related charge, found in Walters (2015, p. 35). If the truth value of the antecedent determines what type of counterfactual a conditional is, then a person would not know the truth condition of the counterfactual she puts forward without knowing the truth value of the antecedent. I think that is correct, but not particularly troublesome. For, there is often a gap between what one *knows* and what one *justifiably believes*. An utterance of such conditionals is well admissible in an average conversational circumstance, as long as the speaker endorses a certain set of background beliefs which are sufficient to prompt the utterance in question. John, for instance, believing that Jupiter's having 76 moons is independent on the truth of the Continuum Hypothesis, may put forward

(6) If the Continuum Hypothesis were true, then Jupiter would have 76 moons.

There is a sense that John is *certain* of the “truth” of (6). The certainty, however, does not amount to a piece of strict knowledge. Rather, it is better seen as an article of justified belief. The justification of the belief is grounded upon the plausibility of his background belief that the number of moons Jupiter has is independent on the truth of certain mathematical conjectures and of his belief that (6) is a proper way of stating that independence. But for the reason laid out earlier in this section, (6) by itself may not really be a way of stating independence. It could equally be a way of stating connection, depending on the truth value of the antecedent, giving that of the consequent. John only *intends* (6) to be a way of stating independence. As a result, John’s belief, though justified, may not be true. Literally, the person in such a situation may not really *know* the content of her utterance, but she certainly knows what she *intends* her utterance to convey.

So the asymmetry charge loses its appeal, which initially seems to be a fairly strong case against classifying true–true counterfactuals as connection-affirming conditionals. Given the plausibility of such considerations, true–true counterfactuals such as (3) finally prove false. And it is solidly established here that the falsity of such conditionals constitute genuine counterexamples to McGlynn’s semantics, as they come out true according to McGlynn’s account when they should be false according to ECH.

Under the ECH considerations, we are in a better position to make it more explicit why McGlynn’s semantics fails in cases such as (3) where Penczek’s succeed. As McGlynn himself notes (2012, p. 281), Penczek’s clause (iii), by requiring us to consider what the truth value of $\Phi > \Psi$ would be if both (true) components were false, is in effect a test of connection between Φ and Ψ. When Φ and Ψ are actually true, there is no obvious way to spot whether their concurrent truth is brought about by a certain connection or it is due to mere coincidence. Penczek’s clause (iii) invites a consideration of this in worlds where Φ and Ψ are false. To speak of closeness in terms of minimal change and validating, if the minimal changes made to the $({\sim}\Phi \,\&{\sim}\Psi)$-worlds to bring about the truth of Φ are sufficient for bringing about the truth of Ψ, then, so the tests goes, Ψ must be connected to Φ in a certain way. What else, anyway, could it be other than the bare fact that a connection obtains between Φ and Ψ, that would best explain why the minimal Φ-validating changes are sufficient for Ψ-validating?

McGlynn’s clause (iii′), on the other hand, cannot serve the purpose of the connection test. For it only requires the consideration of $\Phi > \Psi$ in worlds where Φ, but not Ψ, is false. Even if clause (iii′) is satisfied by some Φ and Ψ which are actually true, it cannot be automatically determined whether the minimum changes required for bringing about the truth of Φ are sufficient for bringing about the truth of Ψ. Anyway, Ψ might be true *because of*, in a very loose sense of the phrase, the truth of Φ, or Ψ might be true merely because it is true *all the way* that what it takes to bring about the truth of Φ leaves the truth of Ψ intact. And it is exactly the latter case when Φ and Ψ are actually true but irrelevant. But it is also exactly such kind of *irrelevance* that underlies the intuitive falsity of such true–true counterfactuals which McGlynn’s account wrongly classified as being true.

4 Conclusion

The problem posed by Conjunction Conditionalization for the standard semantics of counterfactuals has received brilliant treatments. Penczek's remedy of the standard semantics seems almost satisfactory, except for the fact that it runs into trouble with semifactuals. McGlynn's proposal, in turn, proves defective in another way. I argued, appealing to an extended version of the connection hypothesis, that counterfactuals with true but irrelevant components constitute genuine counterexamples to McGlynn's semantics.

In the process, I also argued for the plausibility of an indispensable asymmetry between semifactuals and true–true counterfactuals, namely that true–true counterfactuals affirm certain connections while semifactuals deny certain reverse connections. The asymmetry, if I'm right, is at the root of the defects of both the accounts proposed by Penczek and McGlynn. Building it back into our semantics seems to be a first step towards doing greater justice to an accurate understanding of counterfactuals and circumventing the cunning troubles posed by Conjunction Conditionalization. I suggest, as a preliminary, that a full theory of counterfactuals takes more or less the following form:

$\Phi \;\Box\!\!\rightarrow \Psi$ is true at world i if and only if EITHER

(i) world i is a Φ-world or a $\sim\Psi$-world, and $\Phi \Box\!> \Psi$, OR
(ii) world i is a $(\sim\Phi \;\&\; \Psi)$-world, and $\sim(\Phi \Box\!> \sim \Psi)$,

where $\Phi \Box\!> \Psi$ is the Penczek Conditional, defined in terms of the Lewis Conditional $\Phi > \Psi$. Things get complicated here. Though remaining broadly standard in substance, defining the truth conditions of counterfactuals this way may spoil the very elegance of the Lewisian legacy. If, however, such a move pays more respect for our linguistic endeavor at marking out the structure of the immanent human thoughts, this sacrifice is inevitable and worthwhile.

Acknowledgments I would like to thank an anonymous referee of this journal for very helpful comments and advices on earlier versions of this paper. I'm also grateful to Aidan McGlynn, Una Stojnic, Simon Goldstein for their valuable comments and kind help on various points. This research was funded by a Humanities and Social Sciences Project of the Ministry of Education of China (No. 13YJC720014).

References

Cogburn, J., & Roland, J. (2013). Safety and the true–true problem. *Pacific Philosophical Quarterly, 94*(2), 246–267.
Gundersen, L. (2004). Outline of a new semantics for counterfactuals. *Pacific Philosophical Quarterly, 85*, 1–20.
Lycan, W. (2001). *Real condtionals*. Oxford: Oxford University Press.
McDermott, M. (2007). True antecedents. *Acta Analytica, 22*, 333–335.
McGlynn, A. (2012). The problem of true–true counterfactuals. *Analysis, 72*, 276–285.
Penczek, A. (1997). Counterfactuals with true components. *Erkenntnis, 46*, 79–85.
Pollock, J. (1976). *Sunjunctive reasoning*. Dordrecht: Reidel Publishing Company.

Walters, L. (2009). Morgenbesser's coin and counterfactuals with true components. *Proceedings of the Aristotelian Society, 109*, 365–379.

Walters, L. (2015). Possible worlds semantics and true–true counterfactuals. *Pacific Philosophical Quarterly*. http://eprints.soton.ac.uk/364938/.

Walters, L., & Williams, R. (2013). An argument for conjunction conditionalization. *Review of Symbolic Logic, 6*(4), 573–588.