Commentary on Sean Walsh’s “Indeterminacy of Compatibilist Counterfactuals in Consequentialist Theories of Right”

Jean-Paul Vessel (jvessel@nmsu.edu)
New Mexico State University
2011 Pacific APA Symposium

1. Introduction

Sean Walsh makes several seemingly controversial claims in his essay. In this commentary, I will focus primarily on two of them: (i) that certain subjunctive conditionals (interpreted in the spirit of David Lewis) lack determinate truth values (or are just plain false) and are thus unsuited to do the requisite work of generating moral statuses for actions on an objective consequentialist scheme and (ii) that the use of speakers’ interests in resolving vagueness in certain subjunctive conditionals entails a kind of relativism incompatible with classical objective consequentialist theories. I will present a few objections to Walsh’s first claim before presenting a probabilistic version of objective act consequentialism formulated in terms of objective, subjunctive probabilities (or probabilistic subjunctive conditionals) that appears resistant to Walsh’s objections.

I will conclude by illustrating the importance of utilizing speakers’ interests to resolve certain vagueness features inherent in subjunctive conditionals. I will argue that subjunctive conditionals thought about or put forth in contexts of consequentialist moral theorizing are—in at least a few respects—radically different from other subjunctive conditionals. The speakers’ interests in such contexts—I will argue—invoke a more severe semantics (than ordinary “run-of-the-mill” English subjunctive conditionals) required for consequentialist subjunctive conditionals. If my arguments succeed, Walsh’s concerns regarding speakers’ interests and relativism may lack sufficient motivation.

2. David Lewis’ Semantics for Subjunctive Conditionals and Walsh’s Charge of Metaphysical Indeterminacy

Walsh is concerned that many subjunctive conditionals (when analyzed via Lewis’ semantics) required for consequentialist moral evaluation are “too course grained” to deliver the results that objective consequentialists crave. Explicit formulations of a couple of versions of objective act consequentialism, an investigation of Lewis’ semantics for subjunctive conditionals, and a presentation of Walsh’s concerns appear to be in order.

Consider this version of objective act consequentialism; call it “C”:

C: An alternative, $A$, is morally right iff no alternative to $A$ would produce a higher balance of pleasure over pain that $A$ would (were $A$ performed).

Walsh utilizes an “ice cream” scenario in his argument. Suppose at some specific time and location, you’re faced with the following alternatives:

$a1$: Go out for some ice cream
$a2$: Refrain from going out for some ice cream
According to C, you are morally permitted to perform $a_1$ just in case the performance $a_1$ would produce a balance of pleasure over pain at least as great as would be produced by the performance of $a_2$. So, whether or not $a_1$ is morally permissible according to C is determined—at least in part—by how much pleasure minus pain would be produced were $a_1$ performed. Walsh claims there is no determinate fact of the matter as to how much pleasure minus pain would be performed were $a_1$ performed; thus, C fails to deliver concise moral implications regarding this case (and a plethora of others).

Before proceeding to Walsh’s argument, let’s explore a particular subjunctive conditional relevant to the case and Lewis’ semantic analysis for subjunctive conditionals (or counterfactuals). Assume—as does Walsh—that $a_2$ is actually performed (that our agent refrains from going out for some ice cream). Consider this subjunctive conditional:

SC1: If $a_1$ were performed, then $a_1$ would produce 10 units of pleasure and zero units of pain.

Walsh claims that SC1 (and billions of subjunctive conditionals relevantly similar to it) is false (or perhaps has an indeterminate truth value).

Now consider this semantic analysis of subjunctive conditionals proposed by David Lewis. Lewis’ view is that a would-counterfactual of the form ‘If it were that A, then it would be that C’ is (non-vacuously) true at a world, $i$, just in case the consequent, C, is true at all of the closest accessible antecedent worlds to $i$. (‘Antecedent worlds’ are worlds where the antecedent is true.)

$$ \text{L(SC): } A \supset C \text{ is non-vacuously true at } i \text{ iff } C \text{ is true at all the closest accessible } A-\text{worlds to } i.$$

How close an antecedent world is to a world where a counterfactual is being evaluated is cashed out in terms of comparative overall similarity. Roughly, when evaluating counterfactuals we are interested in what occurs in the antecedent worlds that, given a context, are ‘most similar’ (in relevant respects) to the world where the counterfactual is being evaluated. Furthermore, this view assumes strong centering (thus the ‘SC’ in ‘L(SC)’), which can be captured in the following two claims: (i) no world is more similar to a world $i$ than $i$ is to itself, and (ii) no other world is even as similar to $i$ as $i$ is to itself. The latter and admittedly more questionable claim will play a role in the second portion of this commentary.

Walsh is concerned particularly with the closeness (or similarity) relations among worlds Lewis utilizes in specifying the truth conditions of subjunctive conditionals. At least two factors are responsible for determining how close one world is to another on Lewis’s view: similarity with respect to laws of nature and perfect match of actual fact throughout space-time regions. Assuming determinism, Lewis imagines a closest antecedent world to a world where a subjunctive conditional is being evaluated (the “world of evaluation” for short, usually the actual world) to be one that shares an identical history with the world of evaluation up until a certain point when the

---

1 Lewis (1973a) and (1979). Notice that L(SC) doesn’t quite capture Lewis’s preferred view. L(SC) invokes the Limit Assumption—an assumption that Lewis rejects because it amounts to ignoring the possibility of infinite chains of more and more similar worlds. Invoking the Limit Assumption for simplicity sake is seemingly harmless in this context; it shouldn’t play any role in the upcoming arguments. For Lewis’s argument against the Limit Assumption, see his (1973a): 19-21.

2 Lewis (1979).


4 Lewis (1979): 47f.
tiniest of miracles required to make the antecedent true takes place. This tiny “miracle” is really just the result of a very slight difference between the laws of nature in a closest antecedent world and the world of evaluation. When evaluating some subjunctive conditionals, the miracle required reflects a substantial difference between the laws of nature in the two worlds. But with respect to subjunctive conditionals utilized in the analysis of morally permissible action on a consequentialist scheme, the miracles are tiny. Given the deterministic and compatibilist assumptions in play, the relevant antecedent worlds are accessible to the agent. By pursing some moral alternative, the agent can ensure that such worlds remain live options for actuality. Agents have the ability to see to it that the antecedents of consequentialist subjunctive conditionals are true.

Let’s return to our particular subjunctive conditional and Walsh’s argument.

SC1: If \(a1\) were performed, then \(a1\) would produce 10 units of pleasure and zero units of pain.

According to \(L(SC)\), SC1 is true just in case its consequent (“\(a1\) produces 10 units of pleasure and zero units of pain”) is true at all the closest accessible worlds to the world of evaluation in which the antecedent (“\(a1\) is performed”) is true. Walsh argues that SC1 is false and that other similar subjunctive conditionals (for example, “If \(a1\) were performed, then \(a1\) would produce 9 units of pleasure and 1 unit of pain”) are also false. Walsh’s reason: There is no unique closest antecedent-world to the world of evaluation; instead, there are many such worlds, and there are differences among those worlds regarding precisely how much pleasure and pain are produced by the performance of \(a1\), the going out for some ice cream.

Lewis asks us to imagine nested spheres of worlds around the world of evaluation (on the strongly centered account). In Walsh’s “ice cream” scenario, the agent opts for \(a2\) (doesn’t go out for ice cream), so SC1 is a genuine counterfactual—the antecedent is false in the world of evaluation. Thus, in order to examine whether SC1 is true, we must inspect all of the closest accessible antecedent-worlds to the world of evaluation and see whether the consequent of SC1 is true in all of them. Imagine all such worlds to be identical—perfect matches of actual fact throughout space-time—to the world of evaluation and each other up until the moment of moral deliberation. Then a tiny miracle occurs in each of the closest antecedent-worlds, the smallest possible breach of the laws of nature of the world of evaluation to make it so that our agent pursues \(a1\) (goes out to ice cream) instead of \(a2\). Perhaps something as tiny as a neuron firing in the closest antecedent-worlds that doesn’t fire in the world of evaluation to put any number of small miracles that alter my behavior and determine that I go out for ice cream.” (7) Furthermore, Walsh argues: “There are multiple times and places where multiple types of small miracles that could be inserted into my brain, and these different things can affect longer-term outcomes of the alternative action.” (7)

Walsh argues that there is a plethora of closest accessible antecedent-worlds in our case (each equidistant from the world of evaluation). In those closest antecedent-worlds, different (maximally) minimal breaches in the laws of nature of the world of evaluation required to produce (causally) the decision to go out for ice cream results causally in different amounts of pleasure and pain experienced by subjects in those worlds. If Walsh is right about this, then there are no determinate amounts of pleasure and pain that would be produced were \(a1\) pursued, which would render C incapable of assigning moral statuses to the alternatives in our case.
3. Reflections on Walsh’s Argument and a Probabilistic Version of Objective Consequentialism

It’s unclear to me whether Walsh’s claims about these subjunctive conditionals hold. He appeals to chaos theory and butterfly effects to justify his claims about differences in pleasure and pain that different tiny neural “miracles” in different closest antecedent-worlds produce. I’m somewhat skeptical that such tiny micro-differences between the “miracles” are powerful enough to bring about the different amounts of pleasure and pain (macro-events) that Walsh describes, but I’m pretty ignorant about chaos theory and butterfly effects, so I’ll assume Walsh is right about that stuff.

Granting Walsh’s claims about these conditionals, a natural move might be to appeal to objective, subjunctive probabilities in the construction of an adequate consequentialist theory. Walsh suggests the move before dismissing it:

So, they might argue, while there is not one utility for an alternative option, there is a precise distribution of probabilities for the utilities that may obtain. I disagree, and would argue that the probabilities themselves are vague and course-grained. (9)

It’s unclear to me whether Walsh’s line of argumentation succeeds here. Recall that the “miracles” occurring in these closest antecedent-worlds are (maximally) minimal breaches in the laws of nature of the world of evaluation required to produce (causally) the decision to go out for ice cream. They are identical with respect to how “small” they are. Antecedent-worlds containing “bigger” miracles are farther away from the world of evaluation than are the closest antecedent worlds. I propose that each of these tiniest possible miracles is equally likely to transpire; these closest possible antecedent-worlds are—after all—equidistant from the world of evaluation. If such a proposal rings true, then a seemingly attractive probabilistic form of objective act consequentialism can be constructed.

Even if it were the case that some tiny miracles were more probable—existed in a greater proportion of the closest antecedent worlds—than other equally tiny miracles, I don’t see why these probabilities can’t be identified as clear, determinate, objective subjunctive probabilities.

Let objective, subjunctive conditional probabilities be represented in this way: p(B//A). 5 Adopting the possible worlds framework introduced in Lewis’s semantics for subjunctive conditionals, p(B//A) can roughly be described as a measure of the proportion of the closest accessible antecedent worlds relevant to a particular subjunctive probability statement in which the consequent is true. 6 For those skeptical of objective, subjunctive probabilities, consider this proposal. Let an agent be in an ideal epistemic state relative to a choice situation just in case that agent has complete knowledge concerning all of the circumstances of the situation and complete causal knowledge. The “objective” subjunctive probabilities in one of these troublesome cases might then be identified with the subjunctive probabilities that a perfectly rational agent in an ideal epistemic state relative to the choice situation would assign to the relevant statements. We could feel comfortable treating these probabilities as objective because (i) they are determined entirely

---

5 Nute introduces this symbol in his (1980).
6 For those interested in various semantics for subjunctive probability statements, see, for example, Chapter VIII of Pollock (1976), Giere (1976), and Chapter VI of Nute (1980). Note also that this “ratio of possible worlds construal” of subjunctive, objective probabilities as they are relevant to the construction of objective consequentialist theories is mentioned briefly in Oddie and Menzies (1992).
by purely objective features of the world (for such features are the only things that a fully rational agent in an ideal epistemic state would rely upon in assigning her probabilities), and (ii) given the epistemic status of the agent and the fact that she is perfectly rational, no relevant objective feature of the world would be left out or “mishandled” in the probability assignments. 

Now consider this new theory:

ProbC: An alternative, $A$, is morally right iff no alternative to $A$ has a higher (objective) expected utility than $A$ has.

The (objective) expected act utility of an alternative is the result of multiplying the value of each causal result of the alternative in the closest antecedent-worlds by the objective, subjunctive probability of that result obtaining given that the alternative is performed, and then summing these products. It might be the case that such probabilities and expected utilities are difficult (or even impossible) to know, but that’s an epistemic matter, not a metaphysical one.

(Alternatively, Ben Bradley’s explication of the concept of the expected intrinsic value of a closest world to $w$ can be utilized in the construction of an objective consequentialist theory. Bradley writes:

There are, naturally, an infinite number of worlds that are most similar to a given world. Those worlds might not all have the same intrinsic value. Partition the set of closest worlds according to intrinsic value, so that all worlds with the same intrinsic value fall within the same partition. For each partition, there is a certain probability that a world from that partition would be actualized [were the relevant moral alternative performed]. Multiply that probability by the intrinsic value of a world from that partition; do the same for all the other partitions; sum the products. The sum is what I am calling the ‘expected’ intrinsic value of a closest world.

(Bradley 2005)

For each alternative in a moral choice situation, there is a specific expected intrinsic value of a closest world to the world of evaluation for that alternative. With this concept in hand, a probabilistic version of possibilist consequentialism can be constructed; call it ‘PROB POSS’. Let PROB POSS be the view that among a set of possible alternatives relevant to a moral choice situation, a person, $S$, is morally obligated to do alternative, $P$, at a time, $t$, just in case the expected intrinsic value of a closest $P$-world at $t$ is greater than the expected intrinsic values of each of the closest alternative-to-$P$-worlds at $t$. Formulated this way, PROB POSS turns out to be roughly equivalent to ProbC presented above.)

---

7 To what extent these probabilities are objective is a delicate matter. Following Pollock (1976), a subjunctive probability has been described as a measure of the proportion of the closest accessible antecedent worlds (relative to a particular subjunctive probability statement) in which the consequent is true. To what extent these “objective” subjunctive probabilities are, in fact, objective depends upon the ultimate nature of reality, more specifically: upon whether these proportions are part of the fundamental fabric of reality. If they are, then these probabilities are as objective as they come. If they aren’t, then they still seem to be “objective” enough to suit any “objective” consequentialist’s purposes.
4. The Role of Speakers’ (or Thinkers’) Interests in Securing the Appropriate Semantics for Consequentialist Subjunctive Conditionals (and Subjunctive Probabilities).

Walsh argues that utilizing speakers’ interests to resolve vagueness in consequentialist conditionals results in ethical relativism. I disagree. In fact, I think that such interests are required to secure a more severe semantics for consequentialist subjunctive conditionals than other subjunctive conditionals. Consequentialist conditionals—in a sense—employ more modal force than many other subjunctive conditionals. They demand a weakly-centered analysis. An example should illustrate why.

Unlike most subjunctive conditionals, some consequentialist subjunctive conditionals with true antecedents and true consequents are false. Consider this case. Suppose that a powerful demon (or some very rich person), with hopes of livening up his day, approaches a young gentleman named ‘Sam’ and makes him an interesting offer. Producing a fair coin, the demon asks whether Sam would like to flip it given the following conditions. If Sam flips and the coin lands heads, then the demon will bring about the Good (Pleasure). But if Sam flips and the coin lands tails, then the demon will bring about the Bad (Pain). And finally, if Sam abstains from flipping the coin, then the demon will leave things pretty much the way they are: neither wonderful nor abysmal, rather somewhere in between. Which of the following alternatives should Sam perform on an objective consequentialist moral scheme?

a3: Flip the demon’s coin
a4: Refrain from flipping the demon’s coin

According to our classical theory C, whether Sam should perform a3 or a4 is determined by which (if any) of the two following subjunctive conditionals with identical antecedents but incompatible consequents is true.

Cf1: If Sam were to flip the coin, then it would come up heads.
Cf2: If Sam were to flip the coin, then it would come up tails.

Suppose that Sam in fact opts to pursue a4, refraining from flipping the demon’s coin. By stipulation, the demon thus leaves the world alone. But if Cf1 is true, then Sam has behaved wrongly according to C. Along lines somewhat similar to Walsh’s, I have argued that neither Cf1 nor Cf2 is true.

Two intertwining reasons can be used to explain why neither Cf1 nor Cf2 is true. The fact that the antecedents of the two counterfactuals are extremely underspecified constitutes the first reason. Note that all that is required for the antecedents to be satisfied is that Sam flip the coin. But truth be told, Sam is capable of flipping a coin in any number of ways. Imagine what could factor in the differences between these ways: the direction in which the coin is tossed, the velocity of the toss, the spin action provided by Sam’s thumb and fingers, and so on. Variations in these influencing factors can be used to generate a plethora of different fully specified ways in which the antecedents of our two counterfactuals might be satisfied. And this implies that there are a vast number of different possible worlds—at least one for each fully specified way that Sam might flip a coin—in which Sam flips the demon’s coin. In one possible world, Sam flips it in direction d1,
with velocity \( v_1 \), spin action \( s_1 \), and so on. In another, Sam flips it in direction \( d_2 \), with velocity \( v_2 \), spin action \( s_2 \), . . . . In another, . . . .

Sam lacks the ability to ensure the truth of the consequent of either \( C_{f1} \) or \( C_{f2} \) in close antecedent worlds. Sam simply lacks the requisite dexterity or control to ensure that if he flips the demon’s coin, it will turn up on the side of his choosing. For not only does Sam lack the ability to select and perform a particular fully specified way of satisfying the antecedents of \( C_{f1} \) and \( C_{f2} \), he also lacks the ability to guarantee that some one or other fully specified way whose performance leads to a result of heads (or tails, for that matter) would be realized if he were to flip the demon’s coin.

The similarity relation—whose job it is to determine the closeness of antecedent worlds in the evaluation of counterfactuals—certainly takes agents’ abilities seriously in this kind of case. And given Sam’s inability to ensure that his coin tosses produce the results of his choice, there don’t appear to be any factors that would influence the similarity relation to grant any special priority (or ‘closeness’) to heads-worlds over tails-worlds. Upon contemplation of all the different fully specified ways in which Sam flips the coin in nearby antecedent worlds, those ways resulting in heads turning up won’t have any special priority over ways resulting in tails turning up, and vice versa. Thus, it should be concluded that in the evaluation of the \( C_{f1}-C_{f2} \) pair, some of the closest antecedent worlds are heads-worlds, and others are tails-worlds.

Now suppose the case to be slightly different. Suppose this time that Sam opts to flip the demon’s coin and that it comes up tails. Does that render \( C_{f2} \) above true? So long as we’re using \( C_{f2} \) in efforts to determine what morality might require from a consequentialist perspective, I think not. Consider the following Principle of Normative Invariance:

\[
\text{NI: An alternative’s moral status does not depend upon whether or not it is performed.}^{9}
\]

A general account of subjunctive conditionals such as \( L(SC) \) implies that whether or not an alternative is actually performed plays an important role in whether or not that alternative should or shouldn’t be performed on a consequentialist evaluation. In the first “demon” case, there’s no fact of the matter as to which side of the coin would turn up were Sam to flip the demon’s coin. In the second case, there is—according to \( L(SC) \)—because the consequent of \( C_{f2} \) (the coin comes up tails) is true at the world of evaluation, the closest accessible antecedent-world to that world of evaluation. But, as has been noted, surely a consequentialist conception of rightness shouldn’t exhibit any bias—one way or the other—towards alternatives that just happen to be performed. It can thus be concluded that \( L(SC) \) isn’t general enough to characterize the nature of subjunctive conditionals employed in certain types of ethical reasoning. Lewis’ weakly-centered analysis of subjunctive conditionals is much more suited for this kind of consequentialist work.

\[
L(WC): A \Box \Rightarrow C \text{ is non-vacuously true at } i \text{ iff } C \text{ is true at all the closest accessible } A-\text{worlds to } i \text{ and all worlds, } j, \text{ that are as similar to } i \text{ as } i \text{ is to itself (in relevant respects).}^{10}
\]

\( L(WC) \) renders \( C_{f1} \) and \( C_{f2} \) false regardless of what Sam opts to do. The probabilistic versions of objective consequentialism presented above require a weakly-centered analysis as well.

---

10 Lewis presents a view like this in his 1973a: 29f.
Note that in different contexts with differing speakers’ (or thinkers’) interests perhaps L(SC) provides the correct semantic interpretation. Imagine, for example, betting contexts. Such contexts are radically different from those of moral appraisal. Speakers’ (or thinkers’) intentions play a strong role in selecting the semantics of subjunctive conditionals, particularly the modal strength of subjunctive conditionals in various contexts.

I have also argued (in Vessel 2009) that the similarity relation employed in the analysis of subjunctive conditionals (and subjunctive probabilities) in consequentialist contexts is weaker than the standard similarity relation used in analyses of subjunctive conditionals. We can replace the standard similarity relation with one almost identical to it except that it fails to register differences between worlds due to the future moral failures of agents whose actions are under evaluation. When an agent’s act is under evaluation, her future moral failures and all of the events causally resulting from such failures simply would not register as differences between worlds according to this new “possibilist”-modified version of the standard similarity relation. Again: Speakers’ interests select both the weakly-centered nature and possibilist-modified similarity relations utilized in the analysis of subjunctive conditionals in consequentialist contexts.

Note also that subjunctive conditionals (or probabilities) in consequentialist contexts do not embody the same sort of vagueness with which Lewis is primarily concerned. Consider these two counterfactuals:  

If Caesar had been in command in Korea, he would have used the atom bomb.
If Caesar had been in command in Korea, he would have used catapults.

These counterfactuals are very vague. Lewis argues that in different contexts, either counterfactual is true and that this is possible because more or less importance is placed on varying types of similarity in different contexts. This type of importance is generated by speakers’ interests in conversational contexts.

Consequentialist conditionals are radically different from the “Caesar” conditionals above. As noted earlier, the antecedents of consequentialist subjunctive conditionals (and probabilities) are capable of being made true by the agent of the act under consideration. The antecedent-worlds in such conditionals are accessible to the agent of the act. Consequentialist conditionals are much less vague in important respects than many other subjunctive conditionals. Walsh’s suggestion that utilizing speakers’ interests in resolving counterfactual vagueness entails some sort of ethical relativism might be—to some extent—theoretically unmotivated.

---

11 Lewis (1973a): 66f.
References


