The Probabilistic Nature of Objective Consequentialism

Theorists have consistently maintained that the most plausible forms of objective consequentialism must be probabilistic if and only if indeterminism is true. They claim: If indeterminism is true, then objective probabilities used to map such indeterminacies must be utilized by objective consequentialist moral theories; however, if determinism is true, probabilities play no role in objective consequentialist theorizing. This standard position, however popular, lacks sufficient motivation. Assume determinism to be true and an attempt will be made to show that attractive forms of objective consequentialism must be probabilistic—and not for reasons related to our epistemic limitations either. In this way, perhaps some light will be shed upon the nature of objective consequentialism. Here’s the case.

Consequentialist normative theories can be classified into two groups: subjective and objective. Subjective consequentialist theories might be characterized as those in which an agent’s beliefs concerning the possible consequences of an alternative (including perhaps: their likeliness to obtain should the alternative be performed, their intrinsic value, etc.) play a prominent role in determining that alternative’s normative status. According to objective versions of consequentialism, these beliefs play no such role.

Objective versions can, in turn, be naturally classified into two groups: “subjunctive” and “possibilist”. Subjunctive versions are formulated in terms of would-subjunctive conditionals. Possibilists forgo such conditionals, opting instead to use a
modal accessibility relation and a world-ranking system in the construction of their theories.

Here it will be argued that all extant objective formulations of consequentialism—subjunctive and possibilist alike—fail to deliver the normative implications that the spirit of objective consequentialism requires. The argument rests upon the claim that certain pairs of subjunctive conditionals with identical antecedents and incompatible consequents are such that neither of the pair is true.

In Section 1, a deterministic case is introduced to motivate the objection. In Section 2, the objection is leveled against subjunctive versions of consequentialism. In Section 3, it is argued that possibilist versions are vulnerable to an analogue of the very objection directed against the subjunctivists. In Section 4, an attempt is made to demonstrate that making the problem case indeterministic doesn’t provide any assistance in rectifying matters. In Section 5, the concept of an “objective” subjunctive probability is introduced. And finally, in Section 6, the concept is utilized in the transformation of a subjective version of expected act utility consequentialism into an objective version, one that is capable of dealing with the difficulties posed by the problematic case. Some ways by which the closest thing to a plausible, objective form of consequentialism might be developed are considered at the end.

1. The Case of Sam and the Demon’s Die

For simplicity’s sake, assume that determinism is true.\(^2\) (A consideration of what would result from dropping the deterministic assumption is pursued in §4.) Now suppose that a powerful demon, with hopes of livening up his day, approaches a young gentleman named ‘Sam’ and makes him an interesting offer. Producing a fair, six-sided die, the
demon asks whether Sam would like to toss it given the following conditions. If Sam tosses it and ends up rolling a “1”, then the demon will bring about the Good (Pleasure). But if Sam tosses it and ends up rolling anything but a “1”, then the demon will bring about the Bad (Pain). And finally, if Sam refuses the demon’s offer to toss the die, then the demon will leave the world pretty much the way it is: neither particularly good nor particularly bad. The fate of the world is in Sam’s hands.

Suppose that Sam—our dedicated consequentialist of the objective variety—opts to play it safe, refusing the demon’s offer to toss the die. Now a series of interesting questions might be asked. When Sam chose to do the latter of the following alternatives:

\[ a1: \text{toss the die} \]
\[ a2: \text{refuse to toss the die} \]

did he do the morally right thing from an objective consequentialist perspective? Many will probably be inclined to think that he did. What do subjunctive versions imply about the normative statuses of the alternatives in this case? What about possibilist versions? The answers to these latter questions, it will now be argued, pose serious problems for extant versions of objective consequentialism. They just can’t generate the required results.

2. Leveling the Objection against Subjunctive Versions of Objective Consequentialism

Subjunctive versions of objective consequentialism come in many varieties. Consider here a naïve, sample subjunctive version to illustrate the trouble that our case (and billions of cases like it) brings about for any extant subjunctive form of objective consequentialism.
According to this subjunctive version (call it ‘SUB’), let an alternative, $a$, be morally right just in case $a$’s outcome is such that no alternative to $a$ has a better outcome than $a$ has. Let the outcome of an alternative, $a$, be everything that would result, were $a$ performed.³ Now a subjunctive evaluation of Sam’s case can be provided.

The normative statuses of Sam’s alternatives are determined—according to SUB—by the values of their outcomes. If one alternative has a better outcome than the other, then that alternative is deemed right and the other wrong. If their outcomes are equally valuable, then each alternative is deemed morally permissible.

It’s been stipulated that the outcome of $a_2$ is neither good nor bad. Sam refused the offer to toss the die, and the demon left the world alone. How the outcome of $a_1$ is to be evaluated, on the other hand, seems to be a tricky matter.

The outcome of $a_1$ is everything that would have happened had Sam tossed the die. Much of what would have happened had Sam opted to perform $a_1$ is of little importance: Sam’s hands would have moved in a die-tossing fashion, a die would have flown through the air before bouncing on the ground, etc. The only feature that seems to be of any particular moral importance in the outcome of $a_1$ is which side of the die, if any, faces upward as a result of Sam’s toss. If Sam would have rolled a “1”, then the value of $a_1$’s outcome is greater than the value of $a_2$’s outcome, implying $a_1$ to be morally obligatory by SUB. But if Sam would have rolled anything other than a “1”, then, of course, the value of $a_2$’s outcome is greater than the value of $a_1$’s outcome, implying $a_2$ to be the morally preferable option by SUB. If there is no fact of the matter as to which side would have turned up had Sam tossed the die, then—it will soon be argued—either the value of the outcome of $a_1$ is undefined or the values of the two
alternatives might very well be identical. Either possibility presents insurmountable
difficulties for SUB.\textsuperscript{4}

Whether or not there is a fact of the matter as to which side would have turned up
had Sam tossed the die might be displayed through an investigation of some of the most
relevant counterfactual conditionals to the case. Here are two counterfactuals that seem
to be particularly relevant:

\textbf{Cf1:} If Sam had tossed the die, then he would have rolled a “1”.

\textbf{Cf2:} If Sam had tossed the die, then he would have rolled something other than a
“1”.

If Cf1 is true, then SUB implies \(a_1\) to be morally obligatory. If Cf2 is true, then SUB
implies \(a_2\) to be morally obligatory. But are there any good reasons to believe that either
of these counterfactuals with underspecified antecedents and incompatible consequents is
true? It’s doubtful.

There is a plethora of different fully specified ways in which Sam could have
satisfied the antecedents of these counterfactuals; call these fully specified ways \(f\)-props.
Sam, however, lacks the requisite dexterity to select and then perform any particular \(f\)-prop.
Furthermore, Sam lacks the ability to guarantee that some one or other \(f\)-prop
whose performance leads to a result of a “1” coming up (or a “2” or a “3” or any
particular side, for that matter) would have been realized had he tossed the demon’s die.
In virtue of these facts, there is no fact of the matter as to which side would have turned
up had Sam tossed the die.

This claim can be further illustrated through consideration of the way in which
these counterfactuals might be evaluated via David Lewis’s semantics for subjunctive
conditionals. Lewis’s view is that a *would*-counterfactual of the form \[ \text{If it had been that } A, \text{ then it would have been that } C^1 \text{ is (non-vacuously) true at a world, } i, \text{ just in case the consequent, } C, \text{ is true at all of the closest accessible antecedent worlds to } i. \] (“Antecedent worlds” are worlds where the antecedent is true.)

\[ L: A \rightarrow C \text{ is non-vacuously true at } i \text{ iff } C \text{ is true at all the closest accessible } A-\text{worlds to } i. \]

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.

In order for either Cf1 or Cf2 to be true according to Lewis’s semantics, either all of the closest antecedent worlds under consideration must be “1”-worlds or all them must be other-than-“1”-worlds. But given Sam’s inability to ensure that his die tosses produce the results of his choice, there don’t appear to be any factors indicating that “1”-worlds are more similar to the world of evaluation than other-than-“1”-worlds are, and vice versa. Thus, it seems that some of the closest accessible antecedent worlds are “1”-worlds; others are “2”-worlds; still others are “3”-worlds, and so on. This, of course, prevents either Cf1 or Cf2 from being true on a Lewis-style evaluation.

Notice that this evaluation of Cf1 and Cf2 doesn’t necessarily depend upon Lewis’s semantic account of subjunctive conditionals. Any adequate semantics for counterfactuals should recognize that certain would-counterfactuals with underspecified antecedents are not true. Given that there is no particular fully specified way in which
their antecedents would be satisfied and that Sam lacks the ability to guarantee the results of his die tosses, neither of the counterfactuals appears to be true.8

Someone might be inclined to object at this point that Cf1 and Cf2 aren’t the most relevant subjunctive conditionals to this case. An objector might admit that if determinism is true, then the laws of nature may not, and typically do not, determine the outcomes of unactualized instances of act types. Rather, these outcomes are dependent upon the particular circumstances of the unperformed act token as well as “the” particular unperformed act token itself. Our objector might then conclude that if we appeal to subjunctive conditionals that are relativized to the total situation and “the” specific unperformed act token in that situation, then determinism entails determinate outcomes for the moral alternatives relevant to Sam’s case. But such an objector would be mistaken.

There is no act token that is both more specific than \(aI\) and capable of serving as a moral alternative. Consider the following f-props that might be appealed to as more specific alternatives capable of generating determinate outcomes in worlds where determinism is true:

\(f1: \) Sam tosses the die in direction \(d1\), with velocity \(v1\), spin action \(s1\), . . . .

\(f2: \) Sam tosses the die in direction \(d2\), with velocity \(v2\), spin action \(s2\), . . . .

It seems obvious that the specific way in which a die is tossed plays an important role in determining which side its top will display. Given constant environmental conditions, the direction, velocity, spin (and so on) of the die determine which side turns up. Thus, the performance of different f-props can produce different results. This is noticeably clear in
consideration of the role that such f-props play as antecedents in counterfactual constructions, just as our objector supposes.

Consider the following pair of counterfactuals:

Cf3: If Sam had tossed the die in direction d1, with velocity v1, spin action s1, . . .
then he would have rolled a “1”.

Cf4: If Sam had tossed the die in direction d1, with velocity v1, spin action s1, . . .
then he would have rolled something other than a “1”.

The antecedents of Cf3 and Cf4 are identical f-props. So, again, whether either Cf3 or Cf4 is true depends upon whether or not either the consequent of Cf3 is true in each of the closest antecedent worlds or the consequent of Cf4 is true in each of the closest antecedent worlds. In the case of this Cf3-Cf4 pair, however—unlike that of the Cf1-Cf2 pair—it does seem to be the case that one of the counterfactuals is true. Since f1, which serves as the antecedent for Cf3 and Cf4, is a fully specified way that Sam might toss the die, it seems that there is only one way that the die will turn up should f1 be realized.9 If we are correct in believing that the precise way that a die is tossed (given constant environmental conditions) determines which side turns up, then it seems proper to conclude that either the die displays a “1” in all of the antecedent worlds relevant to the evaluation of the Cf3-Cf4 pair or the die displays something other than a “1” in all such worlds. Thus, either Cf3 or Cf4 appears to be true.

But notice that f1 (the antecedent of Cf3 and Cf4) and f2 are incapable of serving as genuine moral alternatives, rendering Cf3 and Cf4 irrelevant to our case. Our objector’s substitute “alternatives”—f1 and f2—are beyond Sam’s control and thus violate the contrapositive of an “ought implies can” principle:
ALT: If S doesn’t have the ability to select and perform A, then A is not a morally relevant alternative (i.e., an alternative that is a bearer of normative status).

Since f1 and f2 satisfy the antecedent of ALT, they remain ineligible as candidates for bearers of normative status, rendering the fact that either Cf3 or Cf4 is true irrelevant to Sam’s case.

An objector might still, for some unknown reason, demand that either Cf1 or Cf2 is true—that the die would have displayed a “1” had Sam flipped it, or that it would have displayed something other than a “1” had he flipped it. To suggest this, however, is to make the seemingly insupportable ontological claim that there is something in the world that makes it the case that were Sam to toss the die, it would turn up on one particular side rather than another. This claim seems implausible. Remember: (i) that Sam never tosses the die, (ii) that the antecedents of these counterfactuals are underspecified in the sense that there is a multitude of different fully specified ways in which Sam could have tossed the die in order to satisfy the antecedent, and (iii) that Sam lacks the ability to ensure the outcome of the die toss: he doesn’t have the ability to select and perform some f-prop from a specific group of f-props (such as those resulting in a “1” coming up) rather than from another (those resulting in something other than a “1” coming up); instead, he blunders into specific f-props in nearby non-actual worlds where he sees to it that the die is tossed. Some of these worlds are “1”-worlds; others are other-than-“1”-worlds. Therefore, to assume that worlds in which the die displays one particular numeral on its top side rather than another are the closest worlds to Sam’s world is wholly unjustified. Neither of the relevant counterfactuals is true.
The fact that neither Cf1 nor Cf2 is true poses an interesting problem for SUB. Since there is no fact of the matter as to what the demon would have done had Sam tossed the die, then it seems that the value of the outcome of a1 is either indeterminate or perhaps identical in value to that of the outcome of a2. Either possibility renders SUB implausible.

First assume that Cf1 and Cf2 are neither true nor false but indeterminate.\(^{11}\) If it’s indeterminate whether the demon would have brought about the Good (or would have brought about the Bad) had Sam rolled the die, then the value of the outcome of a1 is also indeterminate, and thus undefined.\(^ {12}\) Without a definite value for the outcome of a1, SUB fails to assign normative statuses to the alternatives in the case. In failing to assign such statuses in moral situations, SUB renders itself implausible. If the value of the outcome of a1 is indeterminate, then SUB lacks the complete generality required of plausible normative theories.

Now assume that Cf1 and Cf2 are false.\(^ {13}\) On this assumption, neither of their consequents is contained in the outcome of a1. If this is so, then it may very well be the case that the value of the outcome of a1 is identical to that of the outcome of a2, indicating each alternative to be morally permissible according to SUB. In the outcome of a2, the demon refrains from doing anything to the world. The outcome of a1 is slightly different. On the assumption that Cf1 and Cf2 are false, the demon does do something to the world in the outcome of a1, but it’s not the case that he does something beneficial to the world, nor is it the case that he does something detrimental to the world. Thus, given our assumption, the outcome of a2 seems to be very similar to that of a1 in all morally interesting respects, which suggests that the two outcomes may very well be
identical in value. In the least, there don’t appear to be any good reasons suggesting that one of the outcomes is more valuable than the other. If this is so—as it seems to be—then SUB implies each alternative to be morally right, indicating that Sam didn’t have an interesting moral decision to make. But he did.

Imagine that the demon’s die were 20-sided, or perhaps even 100-sided. It still wouldn’t be the case that something other than a “1” would have turned up had Sam rolled the die. Yet it’s doubtful that anyone sympathetic to an objective form of consequentialism would suggest that this fact renders the normative status of rolling the die identical to the normative status of refraining from rolling. If Sam had performed \( a_1 \), he would have posed a huge, seemingly morally unjustifiable risk to the world. This morally relevant feature of the case just doesn’t seem to be given the required consideration by theories like SUB. Theories like SUB appear incapable of delivering plausible implications (or any implications at all!) when none of the most morally relevant would-counterfactuals to a case is true. Possibilist theories, it will now be argued, fare no better.

3. Leveling the Objection against Possibilist Versions of Objective Consequentialism

Possibilist versions of objective consequentialism were introduced to deal with some independent formal problems plaguing their subjunctive counterparts, most notably: problems involving the specification of relevant alternative sets. In overcoming many such problems, possibilism has proved to be a success. But extant versions of possibilism seem to be no better suited to deal with the problems posed by our case than SUB is. Consideration of a paradigm possibilist view will illustrate why.
Possibilists don’t make use of the problematic would-conditionals used to explicate the concept of an outcome in subjunctive versions of objective consequentialism. Instead, a modal accessibility relation and a world-ranking system are utilized to elucidate the normative status of an alternative. Let’s consider the pioneering, paradigm possibilist theory introduced by Fred Feldman.  

Roughly, Feldman’s view is that a person, S, is morally obligated to do something, P, at a time, t, just in case S does P in all of the best accessible worlds to S at t. A “best” accessible world is one than which no other accessible world is better. The accessibility relation is left undefined as a conceptual primitive; its role is to articulate what an agent can do at a time, that is, which possible worlds an agent can keep open as live options for actuality. Call this rough version of Feldman’s view ‘POSS’.  

POSS, somewhat surprisingly, implies that Sam was morally obligated to roll the demon’s die. Why? Because Sam rolls the demon’s die in all of the best accessible worlds to him at the time. Sam obviously could have rolled the die; thus, worlds were accessible to him (in the relevant sense) in which he in fact rolls it. In each of these accessible “die-rolling” worlds, the top of the die displays a particular numeral. Some of these worlds are “1”-worlds, worlds in which Sam’s toss results in a “1” coming up. Others are “2”-worlds; still others are “3”-worlds, and so on. This is so because in some of the closest accessible die-rolling worlds Sam blunders into an f-prop resulting in a “1” coming up; in others, he blunders into an f-prop resulting in a “2” coming up, and so on. Sam can—accidentally—access a “1”-world; it’s possible that he rolls a “1”. As stipulated by the case, all of the best accessible worlds to Sam at the time were “1”-worlds, worlds in which Sam’s roll of the die leads to the demon’s bringing about the
Good. Since Sam rolls the demon’s die in all of best accessible worlds to him at the time (in all of the “1”-worlds), POSS implies that Sam is morally obligated to roll the die. But surely this implication isn’t consonant with what the spirit of objective consequentialism requires.

Possibilists might reject this conclusion, claiming that despite the preceding argumentation “1”-worlds simply weren’t accessible to Sam at the time. Since Sam doesn’t have the ability to control how the die lands, he can’t “see to it” that he rolls a “1”, suggesting that such a world simply wasn’t accessible to him. And, of course, if “1”-worlds weren’t accessible to Sam, then he did the right thing according to POSS by refusing the demon’s offer. Thus, it may be objected that I have mischaracterized the possibilist view. But I don’t think that I have despite alleged evidence to the contrary.

In my view, to claim “1”-worlds weren’t accessible to Sam is ontologically unsupportable. If this possibilist objection rings true—if possibilists can dodge my objection by claiming that “1”-worlds weren’t accessible to Sam—then Cf2 is true.

Cf2: If Sam had tossed the die, then he would have rolled something other than a “1”.

But I have taken great care to establish that it isn’t. The burden of proof in this case lies upon the person who claims that “1”-worlds are inaccessible—that there’s something in the world that serves as a truthmaker for Cf2. But there is no such truthmaker. Was there something preventing Sam from manipulating his fingers in a way that would result in the die coming up “1”? No. One must have some reason for declaring Cf2 to be true in light of my arguments. It appears to me that there are no good reasons to believe that “1”-worlds weren’t accessible to Sam and that Cf2 is true. It was possible for Sam to
maneuver his fingers in ways that would result in a “1” coming up; “1”-worlds were accessible to him (in Feldman’s sense of accessibility). It’s just that sometimes it’s very difficult to access accessible worlds. This “non-accessibility” defense against my objection appears to fail.

Still, some might continue to consider the claim that “1”-worlds were accessible to Sam to be a disputable one. For those, an alteration of the case should produce the desired results. Forget about the die. Suppose, instead, that the demon tells Sam that he is thinking of a number between 1 and 1,000, inviting Sam to guess which number it is. If Sam declines to guess, the demon promises to the leave the world as it is. If Sam guesses correctly, the demon will bring about the Good; and if Sam guesses incorrectly, the demon will bring about the Bad. In this case, it seems that possibilists will have to admit that there is a world accessible to Sam in which he guesses correctly. This is because every number between 1 and 1,000 is such that Sam can see to it that he guesses this number. Hence, POSS implausibly implies that Sam ought to make a guess.

The problem with POSS’s evaluation of these cases is that it grants no moral relevance to the fact that Sam doesn’t have the requisite ability to ensure that he would have accessed a best accessible world had he rolled the die (or made a guess). He might have accessed a worst accessible world had he taken up the demon’s offer. Our lack of ability to secure certain results prevents us from accessing the best worlds accessible to us in many cases. And this fact must be accommodated by any plausible, objective consequentialist theory. Unfortunately, it isn’t accommodated by POSS. POSS, too, fails to provide plausible implications when none of the most morally relevant would-counterfactuals to a case is true.
4. Making the Case Indeterministic Won’t Help the Consequentialist Cause

One might be led to believe that introducing the notion of an *objective* probability might prove helpful to the objective consequentialist cause. Perhaps if a theory like SUB or POSS were modified so that objective probabilities played essential roles in determining the normative statuses of alternatives, then the problems posed by our case would be overcome. It’s unlikely, however, that introducing such probabilities, as they are usually interpreted, will generate the desired results.

Objective probabilities have traditionally been used to map indeterministic features of the world. In our case, however, the assumption is that determinism is true, implying that there are no indeterministic features relevant to the case that traditional objective probabilities might be used to map. But what if this deterministic assumption were dropped? Would standard objective probabilities then be capable of coming to the rescue? It’s doubtful.

Suppose that there is some kind of quantum indeterminacy in Sam’s world. Then we would be able to map such indeterminacy with objective probabilities. But the problems posed by our modified case would still be very much independent of indeterministic concerns. It’s Sam’s *inability* to secure desired results in die tosses that generates the problems. It’s not as if the demon handed Sam a “quantum” die, one such that indeterministic quantum events would have been solely responsible for preventing Sam from guaranteeing the desired result. Indeterminacies might further complicate matters, but using objective probabilities to map them out leaves the troubles posed by Sam’s inabilities still to be dealt with. *Subjunctive* probabilities (or perhaps probabilistic
subjunctive conditionals), on the other hand, might prove capable of dealing with both of these problems.

5. “Objective” Subjunctive Probabilities

Something like subjunctive probabilities, it seems, must be utilized in order to deal with the problems that Sam’s inabilities pose. The type of subjunctive probability that seems most relevant to our case is a kind of conditional probability. Neither Cf1 nor Cf2 is true; there is no fact of the matter as to what the demon would have done had Sam tossed the die. But there does seem to be a fact of the matter as to what the probability is that the demon would have brought about the Good (or the Bad, for that matter) had Sam rolled the die.

Consider the following sentences.

R: Sam rolls the demon’s die.

G: Sam rolls a “1”, and the demon brings about the Good.

B: Sam rolls something other than a “1”, and the demon brings about the Bad.

What we’re interested in here is the relevant probabilities of G given R and B given R.

To start off, let’s consider these probabilities to be subjective probabilities. Note first of all that these conditional probabilities are not the same as those traditionally utilized in theories of rational belief. The conditional probability of B given A, under the interpretation used in theories of rational belief, is usually defined in this way:

\[ p(B|A) = \text{df. } p(A \& B)/p(A) \]

But obviously this interpretation won’t suit our purposes. According to this interpretation, the likelihood that R will obtain plays an important role in determining the conditional probabilities p(G|R) and p(B|R). But what we’re interested in here is whether
or not Sam should see that R obtains from the perspective of objective consequentialism. The probability that Sam will in fact do something shouldn’t play any role in determining whether or not Sam should or shouldn’t do that something. This is a point that Donald Nute has made clear in his discussion of subjunctive probabilities:

Presumably in this case we are trying to decide whether or not to do such and such. One important consideration is whether B will result from doing such and such. Is the important probability, then, $p(B/\text{we do such and such})$? If so, we must despair, for we surely cannot determine the probability of the antecedent as an intermediate step in making our decision in such a case since what we are trying to decide is whether or not to bring it about that the antecedent is true. . . . Whatever this conditional probability is, it must not be the standard conditional probability of which Adams treats since otherwise it can in no wise enter into our deliberations.

It’s subjunctive conditional probabilities that we’re interested in here, different in kind from the conditional probabilities utilized elsewhere.

To distinguish them from other kinds of conditional probabilities, let subjunctive conditional probabilities be represented in this way: $p(B//A)$. 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. For example, $p(G//R)$ might be thought of as a measure of the
proportion of the closest accessible R-worlds to Sam’s world in which G is true. Now let’s turn to Sam’s subjunctive probability assignments.

What probabilities might Sam have assigned to \( p(G//R) \) and \( p(B//R) \) that led him to refrain from rolling the demon’s die? If Sam were fully rational and had complete knowledge of all of the relevant circumstances—and let’s assume that he was and did—then the probabilities would be \( 1/6 \) and \( 5/6 \), respectively. And these probabilities are perfectly capable of capturing Sam’s inability to guarantee the desired results of his die tosses. But a problem remains. These probabilities are subjective. Objective consequentialist theories provide normative implications without appealing to subjective elements. How might this problem be overcome? Perhaps by showing how these subjective, subjunctive probabilities are determined by objective features of the world.

There is a sense in which it seems that the subjective probabilities Sam assigned to \( p(G//R) \) and \( p(B//R) \) are identical to “objective” probabilities. It’s already been assumed that Sam was fully rational and possessed all the knowledge relevant to his situation. If we gave Sam complete knowledge of his world, he still wouldn’t alter his probability assignments. Given Sam’s epistemic situation and rational disposition, purely objective features of his world—namely, his inabilities—were responsible for his probability assignments. Sam’s probability assignments matched up perfectly with the seemingly “objective” probabilities of \( p(G//R) \) and \( p(B//R) \).

There are billions of cases like Sam’s, cases in which none of the most morally relevant would-counterfactuals is true. Many of these cases are much more complicated than Sam’s. Athletes wonder whether they should attempt feats on the borders of their abilities; snipers wonder whether they should try to secure hits just beyond their ranges in
complicated situations. In all of these cases, it seems that “objective” subjunctive probabilities play some role in determining what is morally required from an objective consequentialist perspective. But unlike Sam’s case, the probabilities assigned by the agents in these choice situations most likely diverge from the relevant objective probabilities, the reason being that they lack huge chunks of knowledge relevant to the case.

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 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.29

With the concept of an objective, subjunctive, conditional probability in hand, various ways in which it might be implemented in the construction of an objective form of consequentialism capable of generating the correct implications in Sam’s case (and others like it) can be explored.
6. Transforming Subjective Versions of Consequentialism into Objective Ones

Decision theorists have long relied upon subjective, subjunctive, conditional probabilities in their theories of instrumental rationality. In recent years consequentialists have borrowed these decision-theoretic devices in the construction of subjective versions of consequentialism. Perhaps the objective version of consequentialism that we are after can be constructed by replacing the subjective, subjunctive probabilities in one of these subjective versions of consequentialism with objective ones.

Frank Jackson has presented a theory capable of the proposed transformation: a subjective version of expected act utility consequentialism. Here’s how Jackson explains how Sam should have ranked his alternatives from the viewpoint of his subjective, consequentialist theory:

The obvious answer is to take a leaf out of decision theory’s book and take the results of multiplying the value of each possible outcome of each contemplated action by [Sam’s] subjective probability of that outcome given that the action is performed, summing these for each action, then designating the action with the greatest sum as what ought be done.30

Jackson isn’t explicit about how these subjunctive, conditional probability statements are to be interpreted. He is sympathetic to the view that they be interpreted as probabilistic conditionals provided that the conditional connective “is read correctly.”31 A weakly-centered version of David Lewis’s semantics for subjunctive conditionals seems capable of providing the correct interpretation. Substituting in objective subjunctive probabilities for Jackson’s subjective ones transforms Jackson’s subjective version of expected act utility consequentialism into an objective version.
How this objective version generates the correct implications in Sam’s case is fairly clear. This new theory implies that an alternative is morally right just in case no alternative to it has a higher expected act utility than it has. The expected act utility of an alternative is the result of multiplying the value of each possible result of the alternative by the objective, subjunctive probability of that result obtaining given that the alternative is performed, and then summing these products.32

Sam’s alternatives were R and ~R.

R: Sam rolls the demon’s die.

There was only one relevant possible result of ~R: the demon leaves the world alone. Let the value of this result be zero. Thus, the expected act utility of ~R is zero.

There were two relevant possible results of R:

G: Sam rolls a “1”, and the demon brings about the Good.

B: Sam rolls something other than a “1”, and the demon brings about the Bad.

In order to determine the expected act utility of R, the values of the possible results of R must be multiplied by the relevant probabilities and then summed. Let the value of G be some very large but finite number, call it ‘n’. Let the value of B be –n. Recall that p(G//R) is 1/6, and p(B//R) is 5/6. The sum, then, representing the expected act utility of R is 1/6n + -5/6n, a number much lower than zero. Sam, by doing ~R, thus did the morally right thing according to this objective version of expected act utility consequentialism.

Note that our sample version of possibilism—POSS—can also be modified in a way that incorporates objective, subjunctive probabilities into its evaluation of behavior.
In his work on virtue consequentialism, Ben Bradley explicates the concept of the expected intrinsic value of a closest world to \( w \). 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.\(^{33}\)

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 possibilism 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 the probabilistic version of SUB presented above. It generates the same results in Sam’s case, and it does so by appealing to the very same objective, subjunctive probabilities outlined above.

Is this objective consequentialist theory then the answer to our consequentialist woes? It handles Sam’s case wonderfully and is perfectly suited to deal with all of the
cases in which agent inabilities render the most morally relevant would-counterfactuals not true.

Unfortunately, while this theory is a step in the right direction, some will still find various features of it unsatisfactory. Theorists attracted to traditional subjunctive formulations of consequentialism might embrace it, but it’s unlikely that possibilists will, for reasons that can only be sketched here. Recall that the theoretical difficulties experienced by those attempting to explicate the essential properties of alternative sets served as prominent motivating factors behind the development of possibilist positions. Paradigm possibilist theories—like POSS—make no appeal to alternatives, nor alternative sets. But PROB POSS does, robbing possibilism of an allegedly, theoretically attractive feature.

Note, however, that there has been no attempt here to expose the complete truth about objective consequentialism; rather, an effort has been made to expose its probabilistic core. And while it’s not quite clear how the most plausible form of objective consequentialism might best be formulated, one thing is for sure: Objective subjunctive probabilities will have to play an essential role in the construction of any plausible version of objective consequentialism to come.
References


1 In 1969, A. N. Prior mapped out this sort of position in the early paragraphs of his “The Consequences of Actions.” For a more recent endorsement of this claim, see the early portions of Howard-Snyder (1997).

2 This assumption shouldn’t play a controversial role in what follows. But some might find the mix of determinism and legitimate moral evaluation repugnant. For them, an additional assumption can be granted; it can be assumed that a tiny “pocket” of indeterminism opens up in an otherwise deterministic world, allowing Sam to perform the alternative of his choice freely. But the assumption doesn’t really seem necessary (especially if you’re a compatibilist). Even if our world is in fact deterministic, providing a semantics for (consequentialist) moral terminology would still be an interesting project.

3 This subjunctive formulation of consequentialism is very similar to the formulation that G. E. Moore presents in his (1912). Philosophers who present objective formulations of consequentialism in this subjunctive manner (with various explications of the concept of an outcome of an alternative) include: Bergström (1966), Bergström (1971), Bergström (1976), Vallentyne (1987), Carlson (1995), and Shaw (1999).

4 Upon hearing this claim, some might find it appalling that an extant, objective, probabilistic version of SUB—or perhaps even some close probabilistic relative of
SUB—hasn’t been appealed to in resolving this matter. For example, Sobel suggests that Jeffrey’s probabilistic account of instrumental (or “expected”) value be employed in the construction of plausible consequentialist principles in his (1971). But the nature of the probabilities that Sobel discusses isn’t perfectly clear. Are they objective or subjective ones? And if they are objective, of which objective variety are they? Furthermore, Sobel’s probabilities are introduced to resolve a puzzle about the outcomes of disjunctions of alternatives, something that isn’t relevant to Sam’s case and those like it. Vallentyne discusses different reasons for introducing probabilistic features into consequentialist theorizing in his (2000), yet similar questions apply. Probabilistic maneuvering similar to that discussed by Sobel and Vallentyne is pursued in upcoming sections.

5 Lewis (1973) and (1979). Notice that L doesn’t quite capture Lewis’s preferred view. L invokes the Limit Assumption—an assumption that Lewis rejects since it amounts to ignoring the possibility of infinite chains of more and more similar worlds. Invoking the Limit Assumption for simplicity’s sake is harmless in this context, for it won’t be playing any role in the upcoming arguments. For Lewis’s argument against the Limit Assumption, see his (1973), pp. 19-21.

6 Lewis (1979).

7 Of course, in other cases it might not only be features about agents’ abilities that play a role in determining the truth values of morally relevant subjunctive conditionals. Fred Feldman has pointed out that the agents’ dispositions to utilize their abilities for good or evil might also play an important role.
For similar evaluations from those who hold semantic views differing from Lewis’s, see Stalnaker (1984), pp. 164-165 and Ayer (1972), p. 116.

It is not being assumed that the set of closest possible worlds in this case is a singleton set. There are, perhaps, a bunch of different ways that Sam can move his body in order to achieve a toss in direction $d_1$, with velocity $v_1$, spin action $s_1$, etc. The point is merely that the worlds in which the fully specified antecedent is satisfied are either all “$1$”-worlds or all other-than-“$1$”-worlds. Clay Splawn is responsible for this point.


Those attracted to Stalnaker’s semantics for subjunctive conditionals might be inclined to evaluate $C_f1$ and $C_f2$ as indeterminate. For a complete account of Stalnaker’s theory, see his (1984), especially ch. 7.

Fred Feldman suggested this SUB evaluation of the case.

Lewis’s semantics suggest this evaluation.

Michael Zimmerman suggested this SUB evaluation of the case.

Contemporary theorists attracted to SUB-versions of objective consequentialism—notably Vallentyne (1987) and Sosa (1993)—have suggested that taking a dangerous risk might be considered to be harmful in itself, and thus intrinsically bad in a way that could affect the values of the outcomes of risk-taking alternatives. In our case, Sam takes a dangerous risk in the outcome of $a1$ whereas he doesn’t in the outcome of $a2$. And perhaps it is this feature that would lead contemporary SUB-theorists to deem $a2$ morally obligatory. But this way of dealing with the case seems to get things backwards. The taking of a dangerous risk doesn’t seem to be bad in itself. If an alternative turns out to be morally wrong in virtue of the fact that it involves taking a dangerous risk, then its
being wrong seems to be determined by what might happen should that alternative be performed. This violates the spirit of SUB, which explains the rightness or wrongness of an alternative in terms of what would happen should that alternative be performed. Theories like SUB just aren’t suited to explain how a1 involves Sam taking a morally unjustifiable risk.

16 See Feldman (1975) and Chapter 1 of Feldman (1986).

17 Also see Zimmerman (1996).

18 See Chapter 2 of Feldman’s (1986) for a precise characterization of the normative theory sketched here.

19 Feldman, in fact, suggests something along these lines on page 24 of his (1986).

20 Subjunctive probabilities seem to play salient roles in efforts to resolve problematic features of certain utilitarian theories in Sobel (1971), Jackson (1991), and Vallentyne (2000).

21 How to make these probabilities as “objective” as possible is on the horizon.

22 There are other problems with interpreting our conditional probabilities in this way. What if the p(R) = 0? Sam may have very well assigned that probability to R. But that would leave the important probabilities p(G/R) and p(B/R) undefined according to this standard interpretation of conditional probabilities. This, in turn, would lend no help to the problems Sam’s inabilities pose for objective forms of consequentialism.


24 Nute introduces this symbol in his (1980).

25 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).

26 It’s not clear how the term ‘proportion’ should be interpreted here. Jake Bridge and Jonathan Schaffer have suggested that in cases in which there is an infinite number of closest accessible worlds (relevant to the evaluation of a particular subjunctive probability statement), the standard interpretation of ‘proportion’ may be unable to secure the desired results, results that are secured in cases involving only a finite number of closest accessible worlds. Let it be left to the mathematicians and ontologists to explicate the relevant interpretation of ‘proportion’ here, an interpretation that may have been merely alluded to in a somewhat metaphorical way.

27 Pollock provides an informal description almost identical to this on pp. 219f of his (1976). Also, a “weakly-centered” interpretation of the phrase ‘closest accessible antecedent world’ is required here, an interpretation that allows for the possibility that worlds other than \( w \) may be as similar to \( w \) as \( w \) is to itself in all respects relevant to the interpretation of a particular subjunctive statement. See my (2003).

28 Others also seem to believe this; see Nute (1980), p. 120.

29 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.


31 Ibid.

32 Mirroring Jackson’s formal account of his subjective version of expected act utility consequentialism, it can be said that an alternative is morally right according to the objective transformation of his theory just in case it maximizes \( \Sigma P(O_i//A_j) \times V(R_i) \), where \( P \) is the objective probability function, \( V \) is consequentialism’s value function, \( R_i \) are the possible results, and \( A_j \) are the possible actions. Jackson (1991), pp. 136-137.

33 See Bradley’s “Virtue Consequentialism,” forthcoming in *Utilitas*.