When Propriety Is Improper*

Kevin Blackwell and Daniel Drucker†

November 21, 2017

Our aim is to clarify the conceptual foundations of the philosophical research program variously referred to by the names 'epistemic decision theory' and 'epistemic consequentialism'. This has been done before,[1] but we think there's profit in doing it again. In particular, we will argue that there are in fact two different projects rather than one, and only confusion can come from conflating them. We'll show that the nearly universally held norm we'll call *Propriety* cannot be true, at least not as a part of epistemic decision theory as we think that project ought to be understood. Since so many results in this burgeoning literature depend on that norm, our aim is unfortunately also destructive. As we'll argue, though, we should have been suspicious all along of the high-powered results that allegedly purely (epistemic) decision-theoretic principles have seemed to deliver. In addition to conceptual clarity, we think that this investigation will also illuminate the substantive normative foundations of any project like epistemic decision theory.

We'll first introduce the essential terminology and conceptual apparatus we and others use. Then we'll examine the main argument in the literature for a requirement related to *Propriety*, which we'll call *Immodesty*. That argument is Lewis (1971)'s *Consumer Reports* analogy. We then show how, given the strongest requirement that that argument can support, extant arguments from *Immodesty* to *Propriety* fail. We then give positive reason to doubt *Propriety*. We conclude by returning to the issue of whether or not the loss of interesting and compelling results proved using *Propriety* should itself cast doubt on what we say; to the contrary, we argue, once you take seriously the

[1]We will mention a few directly as we go, but see also, e.g., Carr (forthcoming) and Schoenfield (2015) for further discussion.

difference between decision theory and ethics, a similar divide between epistemic decision theory and epistemology reveals itself as an attractive way to regard these projects. We should expect epistemic decision theory not to be as substantive as epistemology.

1   Preliminaries and Lewis's Argument

In this section, we'll explain what we take epistemic decision theory to be, and then we'll explicitly formulate the best versions of those claims, specifically in ways that are sensitive to the best ways of arguing for those claims.

First, we'll introduce the jargon. A *credence function Cr* is a function from propositions p in an algebra Ω, closed under complementation and intersection, to real numbers in the interval $[0, 1]$. For simplicity's sake, we'll assume that $\Omega = \mathcal{P}(W)$, where W is the set of possible worlds. We'll also assume, unrealistically, that W is finite. Intuitively, if $Cr(p) = x$, then the person whose credences are represented by Cr is confident to x degree in p.

We'll be interested in credences' *accuracy*, i.e., their distance from the truth. The trouble is that there are a number of ways of measuring distance from $Cr(p)$ to 0 or 1, and the differences between them are crucial. Call a method for measuring that distance a *scoring rule*. We'll use '$\mathcal{I}$' to denote the relevant scoring rule. We could, then, use the absolute value scoring rule: if p is true and $Cr(p) = x$, $\mathcal{I}(x) = |1 - x|$; and if p is false, $\mathcal{I}(x) = |x|$. Or we could use the Brier score: if p is true, $\mathcal{I}(x) = 1 - (1 - x)^2$; and if p is false, $\mathcal{I}(x) = 1 - x^2$. As we suggested, these scoring rules have interestingly different properties, but they'll only be relevant later. What's important now is that you can only compute expected accuracy relative to a way of measuring distance—a scoring rule.

There's another way of conceptualizing these inaccuracy measures, namely as epistemic utility functions. We think that, just as an agent has a generic utility function—a function that determines that, for her, getting a million dollars with no other relevant consequences is better than getting just ten dollars—, agents also place value in believing truths and disvalue in believing falsehoods. Scoring rules encode these values; they are *epistemic utility functions*. Moreover, we think that the standard arguments for the two claims we mentioned in the introduction *require* this interpretation.

That's what we'll try to establish now. Before we do, though, we should emphasize that we understand there are objections to thinking of scoring rules this way; we just think that the philosophers we discuss must conceive them this way, regardless.

The basic thought in epistemic decision theory is this: an agent is rationally required not to φ just when there's some incompatible "act" ψ such that ψ-ing is expectedly epistemically better than φ-ing, where epistemic betterness at least in part depends on the accuracy that results from φ-ing and ψ-ing. We'll limit our epistemic "acts" to transitions between agents' credence functions. So, suppose we have some resultant credence function Cr that we want to evaluate. In that case, the intuitive characterization still leaves two degrees of freedom, since we're talking about expected value.

More precisely, given a set $\mathbf{S}$ of possible states of the world and an agent S's credence function Cr and utility function U, the *expected utility* of φ-ing is $EU(\varphi) = \sum_{s \in \mathbf{S}} U(s \wedge \varphi) \cdot Cr(s)$.[2] Similarly, we can define the *expected inaccuracy* of S's having credence x in p, given that she has a credence of y in $\neg p$. Let '$\mathcal{I}_p$' mean '$\mathcal{I}$ given p'. Then the expected inaccuracy is $EI(Cr(p) = x, \mathcal{I}) = \mathcal{I}_p(x) \cdot x + \mathcal{I}_{\neg p}(x) \cdot y$. We can now also give the following definition:

Definition 1. The *overall expected inaccuracy* of a credence function Cr relative to a scoring rule $\mathcal{I}$, $EI(Cr^*, Cr, \mathcal{I})$, (where '$Cr^*$' names the credences with which we evaluate the potential epistemic act in question, and where 'Cr' names the credences that would result from that potential epistemic act) is given by the following:

$$\sum_{p \in \mathcal{P}(W)} \mathcal{I}_p(Cr(p)) \cdot Cr^*(p) + \mathcal{I}_{\neg p}(Cr(p)) \cdot Cr^*(\neg p).$$

So, first, we need a set of credences with which to calculate the expectation, and we also need an (epistemic) utility function that is at least in part sensitive to accuracy. The only especially natural choices as to those credences and utilities are the agent under evaluation and the theorist doing the evaluating. So, consider the table below:

Agent's credences, Agent's utilities	Agent's credences, Evaluator's utilities
Evaluator's credences, Agent's utilities	Evaluator's credences, Evaluator's utilities

---

[2] Throughout, we use 'φ' to denote both the action, φ-ing, and the proposition that S φ-s interchangeably.

In the top left, we get a very internalist kind of epistemic consequentialism: we calculate expected epistemic value relative to the agent's credences and epistemic utilities. This is the epistemic analogue of standard decision theory. In the top right, we have a less internalist sort. Here, we calculate the expected epistemic value relative to the agent's credences and the theorist's epistemic utilities. In ethics, this corresponds to a special kind of subjective consequentialism: evaluating the agent's actions in light of the agent's own credences, but with utilities as determined by the theorist.[3] The bottom-left corner is too weird to be relevant here (though there are some cases where we evaluate this way). Finally, the bottom-right corner corresponds to a kind of objective consequentialism: it evaluates actions by what are the "right" utilities (at least by the theorist's lights).

Which version of epistemic decision theory you go in for should depend on what you want to use the theory to do or to model. Just as there's no context-neutral answer to whether decision theory or consequentialism is "right", there won't be a context-neutral way of determining which version of epistemic consequentialism is "right", either. But for our own parts, we think that the typical uses to which epistemic decision theory is put suggest the top-left corner, i.e., the analogue of decision theory. That's because we're trying to say when an agent is *rationally required* to do or not to do a given "act", rather than when that "act" would in fact be best. Moreover, the most prominent argument for requiring immodesty seems to depend on the top-left construal of epistemic decision theory. Let us explain. First, we'll give an informal characterization of immodesty, then a more formal version, and finally the argument we just mentioned.

Informally, immodesty says that one ought not to expect (in the sense familiar from statistics) one's own credences to be less accurate than some other credence function. Suppose Cr uniquely accurately represents S's credences. Then S is (credally) *modest* when there's some credence function Cr' such that, according to Cr, Cr' will be strictly more accurate than Cr; otherwise, S is immodest. Slightly more formally:

Definition 2. Let Cr be S's credence function, and $\mathcal{I}$ the epistemic utility function relevant for evaluating S's epistemic rationality. Then S is *immodest* iff there does not exist a credence function Cr' such that $EI(Cr, Cr', \mathcal{I}) < EI(Cr, Cr, \mathcal{I})$; otherwise, S is modest.

That characterizes when an agent *is* immodest, but there is no normativity in that definition. We

---

[3]For this distinction, see, among others, Railton (1984).

want some kind of requirement. How, exactly, the rational requirement to be immodest should be formulated is a matter of (sometimes implicit) controversy. We'll first give the prominent argument for the requirement that we mentioned before, and then the requirement we think it motivates.

In a somewhat different context,[4] Lewis (1971, p. 56) says the following: "Suppose you did trust some non-immodest method. By definition, it estimates some competing method to be more accurate than itself. So if you really did trust your original method, you should take its advice and transfer your trust to one of the competing methods it recommends." We can regiment this argument as follows.[5]

P1. Where φ-ing and ψ-ing are epistemic acts, if the expected epistemic utility to S of S's φ-ing is strictly greater than S's ψ-ing, and it is possible for S to φ, and S's φ-ing is incompatible with S's ψ-ing, then S is (epistemically) rationally required not to ψ.

P2. If Cr is modest given S's epistemic utility function $\mathcal{I}$, then there's Cr' such that the expected epistemic utility to S of S's having credences representable by Cr' is strictly greater than S's having credences representable by Cr.

P3. It is possible for S to adopt credences representable by Cr'.

P4. S's adopting credences representable by Cr' is incompatible with S's having credences representable by Cr.

C. So, if Cr would make S modest, then S is (epistemically) rationally required not to have credences representable by Cr.

Each premise is initially plausible; we'll discuss them in turn.

With P1, we first need to clarify 'epistemic act'. Minimally, it is sufficient for something to be an epistemic act if it's an adjustment or adoption of a degree of confidence in a proposition or propositions. So, for example, making your credence higher that Al Franken will win an election after viewing a debate will count as an epistemic act. That is, strictly, all we need to say for the

---

[4]"Somewhat different" because Lewis is evaluating "inductive methods", i.e., functions from evidence to credences, rather than credence functions in general. We believe that the intuition remains compelling when translated to the present case.

[5]See also Gibbard (2007).

argument to go through; we can leave the rest of the notion intuitive. Another example, however, might prove important for later. Upon considering new evidence, one might *suspend their judgment* in p, where we interpret this as choosing a credence function that is undefined for p.[6]

The justification of P1 is simple. It's a maximizing principle, much like the familiar decision-theoretic injunction to maximize one's expected utility, given a set of choices that one can make. The reasonableness of that principle should make P1 seem reasonable, too. But we would like to emphasize that this *is* a substantive commitment; if you favor some different decision rules, e.g., minimax, you won't find the requirement to be immodest even *prima facie* compelling. Because we think that rules that enjoin maximizing expected utility are good ones, we won't press this any further.

P2 follows from the assumption that S is modest as a result of being representable by Cr.

P3 is perhaps the most obviously controversial of the premises. We have it and the corresponding part of P1 because we know that there are some people who think that the rational 'ought' implies some sort of 'can'.[7] We're less sure that we need it. Even so, though we expect that this is a place some people *will* resist the argument for immodesty, we will not ourselves focus on it.

P4 follows from the fact that an agent cannot have two (different) credence functions at once. Cr and Cr' *must* be different, since they have different expected epistemic utilities, given $\mathcal{I}$. Of course, it might be that one can have *imprecise* credences, where an imprecise credence function is a set of credence functions—but this is just to say that the agent doesn't have (isn't representable by) any precise credence function in particular.

We claim that the rational requirement to be immodest that this argument motivates comes to this:

Immodesty. All agents are rationally forbidden from having credence functions and epistemic utility functions that render that agent modest.

As we conceive of it, this is a wide-scope requirement. It says that a person ought not to have an epistemic utility function and credence function that together conspire to render them immodest.

---

[6]This might not be the right analysis of suspension of judgment. See Friedman (2013) for discussion. For this paper, we're more interested in partial credence functions than suspension of judgment, at least if that analysis turns out to be incorrect; we use the phrase 'suspension of judgment' merely for convenience.

[7]See, e.g., Wedgwood (2013).

They can fix the problem by changing how they value truth and disvalue falsehood, or by changing their credences. Of course, there might be other rational requirements that together entail that, say, the Brier score is the rationally required epistemic utility function. But by itself, the requirement of immodesty imposes no such strong constraints. This contrasts with some other formulations in the literature. Joyce (2009), for example, says that the only admissible epistemic utility function with which to (epistemically) evaluate an agent will be be ones that render that agent immodest in cases where the agent's credences are "clearly rational". This doesn't require that the agent herself have that epistemic utility function. The problem is that if the agent doesn't have the theorist's epistemic utility function, then they won't necessarily be doing worse than they could by their own lights. A version of the Lewis-Gibbard argument that didn't tie the relevant epistemic utility functions to the agent herself wouldn't be plausible. This is because what is so compelling about the Lewis-Gibbard argument is that the agent can see, by her own lights, that she has superior alternatives to her current credence function. It is relatively uncontroversial that the agent's own epistemic utility function is relevant to the evaluation of "acts"; it is far more controversial, and requires *independent* argument, to think that some theorist's epistemic utility function matters for the evaluation of someone else. So, for the Lewis-Gibbard argument to work using a different epistemic utility function than the agent's, we need a separate argument establishing that the theorist's epistemic utility function is normatively relevant for the agent under evaluation. But if the feature that makes the theorist's epistemic utility function normatively relevant is that it makes the agent's credences immodest, the Lewis-Gibbard argument would simply beg the question.

For these reasons, we will assume that agents *have* epistemic utility functions. The Lewis argument shows that at least one very prominent strain of epistemic decision theory conceives of it as an internalist's enterprise, and that enterprise only works at all if agents have epistemic utility functions in the same way that agents have utility functions more generally. In the same way that an agent's desires—both their contents and their strengths—determine, likely not uniquely, a utility function, we would like to say that an agent's desires for believing truths and disbelieving falsehoods determine, likely not uniquely, an epistemic utility function. Their behavior will reveal these utilities, at least in part: with respect to the truth, how risk averse or seeking are these agents? So, going forward, we will assume just the same kind of realism about epistemic utilities that some decision

theorists and ethicists assume about utility functions generally.

Another kind of argument for *Immodesty* is abductive: it shows that we can derive a lot of desirable epistemic principles, so long as we can appeal to *Immodesty*. For example, Joyce (1998) famously argues for probabilism (the requirement that credence functions be probability functions); Greaves and Wallace (2006) and Leitgeb and Pettigrew (2010) argue for conditionalization; and Pettigrew (2016) argues for, among other principles, versions of the Principal Principle and the Principle of Indifference. As we said, we'll say why we don't find this kind of argument compelling in a later section.

Those are the major arguments for *Immodesty*; we'll have more to say about them in a moment. We now want to introduce the second, related claim. But first, we'll need another bit of technology:

Definition 3. An epistemic scoring rule $\mathcal{I}$ is *proper* iff for all possible probability functions Pr and finite sets of possibilities $\mathcal{S}$, and every credence function Cr, $EI(Pr, Pr, \mathcal{I}) \leq EI(Pr, Cr, \mathcal{I})$.[8]

Propriety. All admissible scoring rules for the evaluation of credence functions' rationality are proper.

Though *Propriety* (or strict propriety, see footnote 8) is widely taken to be motivated by *Immodesty*, we think that is wrong. So, consider the following argument from Joyce (2009):

R1. No epistemic utility function $\mathcal{I}$ that is admissible for evaluating the rationality of epistemic "acts" should render any credences modest when there are epistemic circumstances under which those credences are clearly the rational ones to hold.

R2. Something like Lewis (1980)'s *Principal Principle* is correct, i.e., every agent S at t is rationally required to be such that $Cr_S(p|Ch_t(p) = x) = x$, where 'Ch_t' denotes the objective chance function at t.

R3. For any probabilistically coherent credence function Cr, there is some possible chance function at a time, Ch_t and some agent S such that S learns that $Cr = Ch_t$.

---

[8]Most of the authors in this tradition use a stronger notion, *strict* propriety, where '$<$' replaces '$\leq$' in our definition. We think, with Mayo-Wilson and Wheeler (2016), that the Lewis-Gibbard argument clearly does not support strictness, whatever it else it may support, since the existence of other equally good credence functions from an individual's point of view does not undermine the goodness of the individual's own credences.

R4. So, any probabilistically coherent credence function can be clearly the rational one to hold.

C. So, no epistemic utility function should render any person with probabilistically coherent credences modest, i.e., all rationally permissible epistemic utility functions are improper.

It's worth making the argument more intuitive. By definition, only proper epistemic utility functions always render any probabilistically coherent agent immodest when evaluated using those utility functions. The *Principal Principle* makes it rational to have *any* credence function that assigns the same probabilities to the same propositions as the chance function does, when the agent learns what the chance function is. Then, the argument claims, there are no restrictions on what probabilistically coherent credence functions an agent might learn is identical with the chances at the time. So, no probabilistically coherent function ought to be ruled impermissible simply by using a given epistemic utility function. Thus, all epistemic utility functions that are acceptable to use to evaluate credence functions' rationality will be proper.[9]

R1 is a minimal constraint on epistemic utility functions; epistemic consequentialism would simply not be plausible as a theory of epistemic rationality were R1 false. With R2, there's room to quibble about the proper form of the *Principal Principle*, since you might prefer Hall (1994)'s *New Principal Principle*, or something similar. We're confident the choice of the form it takes doesn't much affect whether the argument works. In any case, neither R1 nor R2 are where we will be focusing our attention. R3 is intuitively plausible, since there don't seem to be *a priori* restrictions on what the chances at a given time could be.

Though this argument is, again, *prima facie* plausible, we think it pretty clearly fails. We have three problems. They all target R3.

To see the first problem, note that there are probabilistically coherent credence functions that satisfy the following: for some p, $Cr(p|Ch_t(p) = x) \neq x$. They will not be credence functions that the *Principal Principle* can rationalize; such a credence function cannot possibly be obtained by learning what the objective chance function is for an agent who satisfies the *Principal Principle*. So, they can (and should) be ruled out *a priori*. So, not every coherent credence function can be made clearly rational for some agent by the *Principal Principle*, and so the argument fails.

---

[9]Thanks to an anonymous reviewer for pressing us to clarify the statement of the Joyce argument.

To be clear: we are giving an argument *against R3*, i.e., against the idea that any probabilistic coherence functions might be the clearly rational ones to have. Here's an oversimple example demonstrating the general point: the agent assigns probability 1 to the proposition expressed by '$Ch_t(p = 1)$, but assigns probability 0 to p and 1 to $\neg p$ (and 1 to $p \lor \neg p$, etc.). Assuming p is the only atomic proposition in the agent's algebra, such an agent is probabilistically coherent, but obviously this credence function cannot be made rational by the *Principal Principle*. Of course, it can be ruled as *irrational*—indeed, probably because it so grossly violates the *Principal Principle* itself. But it is a counterexample to R3, because it is a probabilistically coherent credence function that cannot be clearly the rational one to have, or at least cannot be because it is made so by the *Principal Principle*. It shows that R3, and thus the Joycean argument for *Propriety*, fails. We said the example was oversimple, by the way; it's extreme. But because there are no *logical* connections between chance propositions about p and p itself, there will be a good many examples like this, many of which would be less extreme. More generally, we are *not* saying that there are no rational restrictions on what epistemic utility functions an agent can have, or on the ones we can admissibly use in rational evaluation of an agent's credences. So, for all we have said, there might be good reasons to rule out, say, the absolute value epistemic utility function. Here we are just concerned to show that *this* argument for *Propriety* fails, because not every probabilistically coherent function can be rationalized in the way the argument requires.[10]

Our other examples work similarly. So, here's the second problem: there are Moore-like cases. So, take the proposition that there are no doxastic states at t, where t is now. Either there are some objective chances that this proposition can't (now) have (e.g., 1), *or* it can have any objective chance, but—for familiar Cartesian reasons—it wouldn't be rational for any agent to be certain in it. There are also, of course, more direct Moorean analogues, for example that it's raining and I have very low credence that it's raining. Any of these would suffice to give examples of credence functions that are notoriously coherent but irrational to have as one's own. It's important to note that these kinds of Moorean credal assignments are what would be supported by the objective chance function via

---

[10] *Propriety*'s basic thought is that *no probabilistically coherent function* ought to be rationally disincentivized by one's epistemic utility function. Our rejection of this argument, then, amounts to this: that case has not been made. There might be other ways to make it, but this very general way won't work. In section 3 we will provide a general argument for thinking that anything like Joyce's argument here will fail, but here, our only aim was to attack premise R3, and thereby Joyce's argument. We thank an anonymous reviewer for helping us to clarify our response to Joyce's argument.

the *Principal Principle*. There could be facts about your neurophysiology that make it very likely that you will have low credence in it raining in some cases where it is in fact raining. So, once again, there are minimally coherent credence functions that cannot be rationalized in the way the argument requires.

Here's our last problem. It's generally agreed that, where p's truthmakers lie entirely before t, $Ch_t(p) = 0$ or 1. So, $p = <$Abraham Lincoln was assassinated$>$. There is *now* no way to learn that $Ch_t(p) = x$, where x is neither 0 nor 1. So, there is no way for the *Principal Principle* to rationalize assigning non-extremal credence to p for agents like us. More generally, there is no agent for which the *Principal Principle* can rationalize every coherent credence function, since for any agent in worlds at all like ours, there will be propositions whose truthmakers lie entirely in the past.

For all these reasons, we're not optimistic about this argument's chances. But it does highlight some of the more important lessons that we take to be central to why *Propriety* is misguided. In the next section, we'll say what, exactly, we take these to be, so that a simple repair is not likely forthcoming.

We think that this argument (R1, ...) hasn't been put in its most general form. What's really required is some expert deference principle, and that need not be any form of the *Principal Principle*. So, for example, if you treat Dumbledore as a guru, then you ought to satisfy something like this constraint: $Cr_S(p|Cr_D(p) = x) = x$.[11] And if you think Dumbledore could, in principle, have any coherent credence function, then you get an argument with an identical structure. Moreover, our first and third objections wouldn't succeed against this argument. That's not to say it doesn't have its own problems. First, the second objection still very much applies. Second, a version of the first worry still applies, e.g., your learning that Dumbledore assigns x to p cannot rationalize the coherent credence function where you thereby assign $y \neq x$ to p. Also, by being an epistemic agent, rather than an abstraction, it becomes less plausible that the guru can justifiably have conditional credences that violate certain evidentiary norms related to, for instance, induction. So, while we think the generalization is worth pointing out and avoids some versions of some of these problems, it doesn't ultimately help.

---

[11] For discussion of this sort of terminology, see Elga (2007).

So far our arguments have been pretty narrow, focusing only on rebutting Joyce's argument for *Propriety*. It's worth worrying whether there is, nevertheless, a more general reason to go in for *Propriety*: it is a way of guaranteeing that agents always satisfy *Immodesty*. But importantly, even if an epistemic utility function is improper, that doesn't mean that any and every credence function an agent with that epistemic utility function has will render her modest. So there are other ways of complying with *Immodesty* than by having a proper epistemic utility function. There is no simple route even from *Immodesty* to *Propriety*. And, as we argued just now, there doesn't seem to be a complex one, either.

2   Rational Mistakes

Mistakes in an agent's past can cause trouble down the line, both for the agent who has made the mistakes and for the theorist trying to give recommendations for the agent's beliefs. Suppose our agent S has an update policy U, and in response to experience $\mathcal{E}$ fails to update her past credences Cr_S to the credences recommended by U, Cr_S^*, instead adopting Cr_S'. This might happen as the result of a "glitch" in S's cognitive functioning, or of some sort of doxastic or epistemic akrasia—endorsing a doxastic or epistemic "act" but doing something else incompatible. Suppose S is apprised of this—she realizes she failed to update by U. It seems it is permissible to simply switch from Cr_S' to Cr_S^*. *Propriety* does not allow this. Thus, *Propriety* is false. That is the argument we will defend in this section.

In a little more detail, here it is:

S1. Suppose S has Cr_S at t and gets total evidence $\mathcal{E}$ at $t' > t$, but fails to use the rationally permissible update policy U that they endorse, and so has Cr_S' at t'; S uses some other updating procedure U'.

S2. Suppose Cr_S^* is what S's credences *would have* been, had S used U rather than U'.

S3. For all admissible $\mathcal{I}$, at t', $EI_S(Cr_S', Cr_S^*, \mathcal{I}) > EI_S(Cr_S', Cr_S', \mathcal{I})$. (By *Propriety*.)

S4. So, it's irrational for S to switch at $t'' > t'$ from Cr_S' to Cr_S^*. (By epistemic utility theory.)

S5. But, it *is* rational for S to switch at t'' from Cr_S' to Cr_S^*.

C. So, *reductio*, i.e., S3 is false.

There are important objections to consider to this argument.

The first points out that S3 uses a *strict* version of *Propriety* rather than the version we've mostly been working with elsewhere in the paper. But note, first, that many authors use the strict version, too. And either way, it would be a massive and unexplained coincidence if the relevant expected inaccuracies were *always* equal. So we just ask you to restrict your attention to the cases in which S3 is true, given S1 and S2.

According to the second objection, agents who realize they failed to update according to their endorsed update policy U and are thus willing to switch their credences to what they would have had had they used U don't really have the credences we stipulated them to have.[12] After all, they don't endorse them—that's why they're willing to switch!

Not endorsing credences as the right ones to have does not ensure that you don't *have* those credences. Epistemic akrasia is possible, e.g., cases in which one believes that one shouldnt believe p, but nevertheless does.[13] We can be alienated from our credences in some sense. Compare someone who is drunk who uses another person's opinion about whether it's safe for them to drive home—it's *because* the person doesn't trust their high credence that it would be safe that they use the other person's credence. They might even say, "well, I still think I could drive home safely, but if you're so adamant about it...". The situation we're describing is not very unusual. There's a lot more to credence's functional role than endorsement, or even bets, broadly construed—there's also phenomenology, verbal expression of that credence, etc. Sometimes we correctly attribute credences on the basis of those other factors.

The third objection is easiest to understand if U is standard conditionalization, i.e., the update rule according to which, when their posterior credences are Cr and between t and t', $\mathcal{E}$ is the strongest thing the agent learns, then their posterior credences Cr' must be such that, for all p, $Cr(p|E) = Cr'(p)$. Now, suppose the agent uses U' rather than conditionalization (for whatever reason). If it's rational to switch upon realizing that, then don't they have *new* evidence that they

---

[13]That doesn't mean it's *rational* to have such credences. See Horowitz (2014), Greco (2014), and Lasonen-Aarnio (2015) for some discussion. We note, though, that even for those those who think it is irrational, surely it is so only after some unspecified and unspecifiable period of time in which one can correct those beliefs.

can thereby conditionalize on? That would not violate *Propriety*.[14] This objection would work if every agent were rationally required to have credences *prior to the realization* that $Cr'(p|cr^*_{\langle t,t'\rangle}(p) = x) = x$, where $cr^*_{\langle t,t'\rangle}$ non-rigidly denotes the credence function that they would have at t' had they conditionalized on the strongest thing they learned between t and t'. The agent must be required to have these conditional credences prior to the realization rather than after it, since otherwise changing the conditional credences to reflect that would *again* violate *Propriety*. So this objection requires that agents who endorse conditionalization must have those conditional credences.

The trouble is we cannot see a motivation for this requirement. It does not follow from the thought that conditionalization is the uniquely rationally acceptable doxastic update policy. More importantly, an agent might never have considered the situation in which they fail to conditionalize properly. If they *do* find that they failed to conditionalize properly, then perhaps they are then rationally required to have those conditional credences and then conditionalize on them, all at once as it were. That won't help, for the reason gave in the previous paragraph. So requiring these conditional credences requires agents to foresee situations that they needn't be rationally required to foresee.

According to the fourth objection, *Propriety* is an idealization. Based on what we've said so far, this can be fleshed out in two different ways. First, we can say that it is only meant to apply to agents who have not made past epistemic mistakes. In that case, it's easy to see how our argument wouldn't apply. Second, we can say that agents who endorse conditionalization *are* required to have those prior conditional credences, since they are more ideal than ordinary agents, and so we can assume certain powers of foresight of them that we can't of normal agents.

Of course, we have no problem with idealization as such. We do, however, think that these kinds of idealizations would significantly narrow the range of cases to which epistemic decision theory or consequentialism is meant to apply. Insofar as the results they can offer are theoretically desirable rational requirements, we would like to be able to say that they rationally constrain normal agents, too. So idealizing is not without cost. More to the point, though, we think there are problems with *Propriety* even as it applies to pretty strongly idealized agents. In the next section, we'll explain.

---

[14]Thanks to [removed] for raising this worry.

3 Epistemic and Practical Decision Theory

Propriety becomes especially implausible when you start thinking about what its analogue for practical decision theory would be. Doing so will also reveal the extreme idealizations that would be necessary to save it.

By "practical decision theory", we just mean the line of decision theories descended from Ramsey, embodied in Savage (1972), Jeffrey (1965), and Joyce (1999). Unsurprisingly it resembles epistemic decision theory in its basic ingredients: it uses the agent's credences and utilities to form expected values of various options, and then it recommends agents select only top-ranked options. The details and intramural conflicts between various ways to develop the theories aren't important for us. First, consider the following definition:

Definition 4. A utility function U is *practically proper* iff for each possible intention I to φ that S might have, for each probabilistically coherent credence function Cr_S, and for each of S's options ψ, $EV_{U,Cr_S}(\varphi) \geq EV_{U,Cr_S}(\psi)$, where EV is the expected value function of the correct decision theory.

According to **Definition 4**, a utility function is practically proper just when, no matter what intention the agent adopts, that intention could not help but maximize expected utility so long as they also have a coherent credence function. So, a utility function is not proper merely if, relative to it, all *actual* intentions an agent has maximize expected utility; rather, as we said, it must make it that all possible intentions an agent might have do so. It is, in other words, simply the practical analogue of **Definition 3**. With that said, consider now how implausible the following practical analogue of *Propriety* is:

Practical Propriety. All admissible utility functions for the evaluation of intentions' rationality are practically proper.[15]

We think *Practical Propriety* is quite mistaken, and the reason is clear: even people with probabilistic credence functions would sometimes fail to maximize in light of those credence functions, given

---

[15]Thanks to a reviewer for pressing us to be clearer on exactly what *Practical Propriety* comes to.

utilities that are mandated by their intrinsic desires.[16] One reason, of course, is computational: we don't expect agents to be good at calculating the very complex magnitudes involved.[17] But such computational limits are really not the point we wish to raise, since it's pretty natural to idealize them away. The problem is more fundamental: we wish to use practical decision theory's rational injunctions in part to *criticize* those who do not comply with it by maximizing expected utility, and *enjoin* those who do not currently intend what would maximize their expected utility to do so. It's true that there's a traditional way of thinking about those decision theories that also runs through Ramsey, Savage, and Jeffrey, of thinking that there is no independent, or no theoretically important independent, reality to an agent's credences and utilities. But this genre of interpretation is unsatisfying for just the reason we just gave.[18] More basically, we just *do* have intrinsic desires and confidences. They might not have quite the structure we'd like credence or utility functions to model, but that is a different claim entirely. We'll assume, then, that it is *possible* for agents to have credences, utilities, and intentions such that the intention the agent actually forms is of less expected value than some other act they know they could intend. That by itself contravenes *Practical Propriety*. The most fundamental reason is, as we said, that we would like to criticize agents who do not maximize expected utility (and, so, those who do not minimize expected inaccuracy). But how might this happen for actual agents, or even relatively, but not completely, idealized agents?

One factor is sheer laziness. This is distinct from computational intractability. Suppose there are just two options, A and B, which differ only in that A is F and B is $\neg F$. And suppose the relevant agent knows their utilities (and, if you like, their credences) perfectly well. They might still decide that it's not worth the trouble to compute the expected value of A as opposed to B. Then they might intend A, even though B in fact has more expected utility. You might worry that the options are not properly described: really they're A with calculation, B with calculation, A without calculation, and B without calculation; and either A or B without calculation is clearly superior to A or B *with* calculation—though how much would, of course, be impossible to know in a way that

---

[16]It's important that we are criticizing *Practical Propriety* and not a practical analogue of *Immodesty*. The reason is that a practical analogue of *Immodesty* is just the normative *requirement* to maximize expected utility. We think that might be a correct normative principle. Our issue with *Practical Propriety* is that we don't think it's a functional constraint on admissible utility functions that they necessarily render every probabilistically coherent agent practically immodest. Thanks to a reviewer for pressing us here.

[17]See, e.g., Feldman (2006).

[18]For more discussion of these points, see Meacham and Weisberg (2011) and Easwaran (2014).

retains genuine optionality. Perhaps, though we think people often forego calculation where it is expectedly more beneficial than not calculating. This might be a sort of akrasia; if so, it is a very common kind. Indeed, akrasia of a more general kind can cause (or constitute) our intending a less good outcome in expectation.

Not *just* akrasia does that, though. Agents sometimes accept decision rules besides ones that enjoin them to maximize expected utility. This choice itself might be irrational,[19] and it might cause them to make subsequent irrational choices down the line. But surely it is *possible*. So if we would like to offer rational criticism of those people for these reasons, then we should not accept *Practical Propriety*. It should be possible to criticize an agent for failing to maximize expected utility relative to their actual utilities, while allowing that those utilities are rationally acceptable.

Finally, an agent might simply not know what their utilities are. For example, you might get excellent evidence that you intrinsically desire something: for example, knowledge of number theory. But then you try it, and you realize you never actually wanted it; you thought you did, because people you admired wanted it, and you thought, and perhaps hoped, you were like them, as you knew you were in many other respects. Failure to know our own desires is unfortunately common, and not at all always a mistake of rationality: our evidence is sometimes misleading, just as evidence is about anything else. So, we would misevaluate people if we required that rational agents' utility functions always have them maximizing expected utility. Now, perhaps this is a problem with practical decision theory; perhaps it'd be more reasonable to go with a kind of "expected expected utility" theory.[20] We remain neutral. What's important is that this is just yet another way a person might fail to maximize expected utility theory, even with an admissible utility function. And it seems like idealizing this way is a *very large idealization*; why shouldn't rationally ideal agents be able to be misled about what they want?

So, *Practical Propriety* is untenable, since we would like to criticize lazy agents, akratic agents, and perhaps also agents that accept incorrect decision rules; and we would like it to be possible for agents with acceptable utility functions to not maximize expected utility because they are misled as to what those utility functions are. The important point here is that analogues of all of these will apply in the epistemic case, too. Epistemic agents can be lazy, akratic, accept bad epistemological

---

[19]For the thought that it is irrational, see Titelbaum (2015).

[20]See, e.g., Boutilier (2003).

principles, and fail to know their epistemic utility functions. In other words, exactly the same reasons for denying *Practical Propriety* should lead us to reject *Propriety*. Epistemic decision theory is not all that structurally different from practical decision theory; its concerns are simply limited to a very specific kind of utility, and very specific kinds of "acts" (epistemic ones).

Here, then, is the structure of our argument. *Immodesty* and *Propriety* typically have internalist justifications, much like principles of decision theory. That means that we should give to epistemic utility functions roughly the same amount of psychological reality as we do to more general utility functions of the kind used in practical decision theory. But when we do this, we find that *Propriety* is wrong for the same reasons *Practical Propriety* is wrong: agents can genuinely have a given rationally admissible epistemic utility function and still fail to maximize expected utility (or minimize expected inaccuracy). The internalist motivations such principles must receive falsify those principles.

4   Practical Decision Theory and Ethics, Epistemic Decision Theory and Epistemology

We think there is a more general lesson to be learned from this discussion. We said in our introduction that philosophers have used 'epistemic decision theory' and 'epistemic consequentialism' to refer to *the same project*. But clearly 'decision theory' and 'consequentialism' refer to quite different projects. A person might fully comply with decision theory's recommendations while being a *morally terrible person*. This is not so with any correct ethical theory.[21] Decision theory is, by design, much less substantive than ethical theory. This is what gives rise to its generally internalist character, since its recommendations are meant to make sense given arbitrary, or almost arbitrary, points of view; not so for consequentialism. A similar distinction, we think, captures the difference between epistemic decision theory and epistemology, at least as we propose now to use those terms.

Here's a different way to put these points. It's no problem for decision theory if it recommends that a serial killer kill; it very much would be an at least *prima facie* problem if ethical theories delivered that result in normal cases. The point is just that decision theory is radically internalist: it takes credence functions and utility functions for granted—*modulo* certain consistency require-

---

[21]Thanks to a reviewer for getting us to clarify what we meant here.

ments, such as, perhaps, transitivity of preference[22]—and then gives recommendations on that basis. Ethics has some further jobs to do. It can, for example, filter out unacceptable utility functions. Or, it can make some options ineligible independent of what the expected utility calculations recommend, even given otherwise acceptable utility functions. For an example of the kind, it might be that deriving a lot of utility from the pain of others is ethically impermissible. If, on the other hand, you think mental states like desire aren't the province of ethics, then you can say that regardless of whether the utilities themselves are acceptable or not, causing others pain for no other reason is impermissible. Importantly, though, neither is the job of decision theory, which is neutral about what the options are and what the good utility functions are.[23]

There's an analogous distinction between epistemic decision theory and epistemology, we think. The burden of this section will be to convince you of that. As we said, decision theory is, at its very heart, internalist. This means that its recommendations have to be something that speak to the agents themselves, as their points of view in fact work. If a theorist thinks some option would be better for them, but if there's no route from the agent's actual coherent credences and utilities to that option, then the theorist's view is of no consequence as far as decision theory goes. We should emphasize that we are not committing ourselves to reasons internalism,[24] or even, necessarily, internalism about *rationality*. This is just how we think the project of decision theory ought to be delineated. And at first glance, epistemic decision theory seems to respect this, at least more than any other epistemic theory that we?ve seen, at least as traditionally understood. For one, its recommendations are formed by taking expectations of credences and (epistemic) utilities. That is a real and important similarity. Also, the Lewis argument has a decidedly internalist feel. But, as we've seen, the theory is not internalist *enough*. And there are different ways in which it isn't internalist enough, depending on certain choices the theorist might make. Some epistemic decision theorists have imposed on individuals epistemic utilities that, not only do they not have, but they seem not to be rationally required to have, either—such as the proper utility functions. But as we saw, even that's

---

[22]More on this as it applies to our argument against thinking of epistemic decision theory as epistemology in a moment.

[23]We are aware of some people who try to derive ethics from decision theory (or at least from game theory), like Gauthier (1986), but we are convinced that that cannot be done. If you do think it can, however, you'll be far less interested in the distinction we make in this section, but we think sufficiently many people share our suspicion to motivate what we'll say.

[24]See Williams (1981).

not enough, because the agent can fail to be aware of their own epistemic utilities, even in broad outline. Internalist motivations cannot deliver anything like *Propriety*. And they shouldn't, because nothing like *Propriety* is true in the practical case: it would be unreasonable to assume even relatively ideal agents always succeed in matching their intentions to the theory's recommendations.[25] So, even if *Propriety* can allow the epistemic decision theorist to make the intuitively correct epistemic criticisms of agents who, say, have non-probabilistic credence functions, we just don't think that's the theory's job, any more than we think decision theory should allow us to criticize coherent killers and sadists.

Of course, coherence matters even in decision theory, you might think. Money pump arguments certainly seem and are often taken to establish that agents ought to have transitive preferences. We suspect that really these are motivated as modeling assumptions, or as adequacy constraints on a theory that can actually recommend behavior, but either way, our problem is not with probabilism, even as a constraint on input credences to an epistemic decision theory. Rather, our problem is with using internalist arguments to motivate a derivation of probabilism with *Propriety*. So, we're not against coherence constraints, even the ones that epistemic decision theorists have argued for; we just think their arguments are not motivated by epistemic decision theory, but if the constraints are good, they need independent motivation.

So, what's left of epistemic decision theory without *Propriety*? Well, what's left is just the epistemic analogue of decision theory. It won't give as many structural constraints as we might have hoped for, but we didn't hope for similar constraints in the practical case for good reason: decision theory is very much a garbage in, garbage out kind of a theory. We're advocating seeing epistemic decision theory in the same way; we think that's what the project's motivations deliver, and that anything else is too substantive. This means that epistemic decision theory is not, as some have seemed to hope, a replacement for epistemology, any more than practical decision theory is a replacement for ethics. But just as many doubt the derivation of ethics from decision or game theory,

---

[25]A reviewer objects: this looks like a criticism of a practical analogue to *Immodesty* rather than of *Practical Propriety*, because it looks like our criticism is of agents' actual choices rather than their utility functions. But this appearance is a little misleading. Our concern is that it be *possible* to criticize agents for failing to maximize expected utility, which is made impossible by requiring for the evaluation of agents' intentions practically proper utility functions. So our criticism is, in the first instance, of a requirement on *utility functions*, though it is true that that criticism derives from our wish to criticize intentions or choices. Recall from footnote 14 that we think the practical analogue of *Immodesty*, conceived of as a rational or normative requirement on (now) credence-utility-intention triples, might very well be right. We thank the reviewer for this objection.

we doubt the derivation of epistemology from decision theory. We wish to emphasize that we don't think we should simply take every agent's actual epistemic utility function for granted—there are almost certainly substantive, rational requirements on those utility functions that a person might fail to conform to. But, we do think that these constraints ought somehow to be internalistically motivated. We take no issue with the part of epistemic decision theory that meets those demands. In particular, there might be rational constraints on epistemic utility functions that correspond to transitivity and independence in practical decision theory. We thus do not reject the epistemic decision theorists' approach to accuracy measures in every particular; our target is the extremely popular *Propriety*. In epistemology, by contrast, it is an open and much debated question whether epistemic constraints must be internalistically motivated.

There are other advantages to our way of thinking. For example, we think that this way of thinking about things allows us to avoid some problems that have proved very difficult for epistemic decision theorists recently. First, there's the notorious problem of epistemic bribes.[26] So, take the following case:

Converted scientist. Clara is an atheist scientist who seeks a grant from a religious organization. The grant agency will only give grants to fund Christians' research, and Clara knows that she wouldn't be able to fool them. Despite the fact that she's thought long and hard about the issue and knows that, on balance, the evidence strongly supports atheism over Christianity, she also knows that this will cause her belief set to be filled with more accurate beliefs than she had before that.

The epistemic decision theorist who aspires to epistemology will have a hard time with this case, because she will recommend that Clara believe in the Christian God, which is standardly taken to be the wrong result. But our epistemic decision theorist need say no such thing; she can say that, although the true epistemological theory would say that in her evidential position Clara ought not to believe in God, she would get more of what she epistemically wants by doing so. So *Converted scientist* would no longer be a counterexample, nor would the rest of the epistemic bribe cases. This is similar to how an ethical theory might have deontological restrictions that would be out of place in a decision theory.

---

[26]See, e.g., Firth (1998), Berker (2013), Greaves (2013), and Konek and Levinstein (forthcoming).

More generally, the epistemic decision theorist need not worry about attacks based on evidentialism, namely the theory that an agent epistemically ought to fit their credences to the evidence, however that's to be spelled out exactly.[27] So, it would avoid problems based on evidentialism that already exist in the literature even assuming *Propriety*,[28] and problems with evidentialism that arise from having improper epistemic utility functions. That is, when an agent has an improper epistemic utility function, our version of epistemic decision theory will, in some cases depending on the exact function, recommend changing credences but not in response to any new evidence. So, for example, the absolute value epistemic utility function recommends extremal credences in all propositions to which one does not initially give 0.5 credence. But this is again no problem if switching credences in that way is bad by epistemology's lights, even if it secures more of the agent's own epistemic good, at least by their own lights.

Finally, the epistemic decision theorist can give no reason to rule out the following probabilistically coherent priors: .99 to some contingent p conditional on every proposition q. We're not absolutely positive epistemology should rule those out, but we do think epistemic decision theory can't, and so anyone who thinks that epistemology can be derived exclusively by using epistemic decision theory cannot rule those out as epistemically bad. This is just the same as the situations we began by describing: decision theory can't tell agents not to prefer not killing, but ethics might be able to. So, we have yet another reason to distinguish epistemic decision theory and epistemology.

We suspect there are many other reasons, but we think we've made our case. The internalist motivations for epistemic decision theory fit terribly with what it has been used for so far. Moreover, the analogy with practical decision theory should lead us to expect that there is a difference between epistemic decision theory and epistemology. Finally, we can avoid or dissolve some problems that have seemed intractable for epistemic decision theory as it's been practiced. That is, once again, not to say that the results the epistemic decision theorist has argued for cannot be right, only that they ought not to be argued for in that way.

---

[27] See, e.g., Feldman and Conee (1985).
[28] See, e.g., Easwaran and Fitelson (2012).

References

Berker, Selim. 2013. "Epistemic Teleology and the Separateness of Propositions." *Philosophical Review* 122:337–393.

Boutilier, Craig. 2003. "On the Foundations of Expected Expected Utility." In *Proceedings of the 18th International Joint Conference on Artificial Intelligence*, IJCAI'03, 285–290. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc.

Carr, Jennifer. forthcoming. "Epistemic Utility Theory and the Aim of Belief." *Philosophy and Phenomenological Research* .

Easwaran, Kenny. 2014. "Decision Theory without Representation Theorems." *Philosophers' Imprint* 14:1–30.

Easwaran, Kenny and Fitelson, Branden. 2012. "An 'Evidentialist' Worry About Joyce's Argument for Probabilism." *Dialectica* 66:425–433.

Elga, Adam. 2007. "Reflection and Disagreement." *Noûs* 41:478–502.

Feldman, Fred. 2006. "Actual Utility, the Objection from Impracticality, and the Move to Expected Utility." *Philosophical Studies* 129:49–79.

Feldman, Richard and Conee, Earl. 1985. "Evidentialism." *Philosophical Studies* 48:15–34.

Firth, Roderick. 1998. *In Defense of Radical Empiricism: Essays and Lectures*. Lanham, MD: Rowman & Littlefield.

Friedman, Jane. 2013. "Rational Agnosticism and Degrees of Belief." *Oxford Studies in Epistemology* 4.

Gauthier, David. 1986. *Morals by Agreement*. Oxford, UK: Oxford.

Gibbard, Allan. 2007. "Rational Credence and the Value of Truth." In Tamar Szabò Gendler and John Hawthorne (eds.), *Oxford Studies in Epistemology, Volume 2*, 143–64. Oxford, UK: Oxford.

Greaves, Hilary. 2013. "Epistemic Decision Theory." *Mind* 122:915–952.

Greaves, Hilary and Wallace, David. 2006. "Justifying Conditionalization: Conditionalization Maximizes Expected Epistemic Utility." *Mind* 115:607–632.

Greco, Daniel. 2014. "A Puzzle About Epistemic Akrasia." *Philosophical Studies* 167:201–219.

Hall, Ned. 1994. "Correcting the Guide to Objective Chance." *Mind* 103:505–518.

Horowitz, Sophie. 2014. "Epistemic Akrasia." *Noûs* 48:718–744.

Jeffrey, Richard. 1965. *The Logic of Decision*. New York: McGraw-Hill.

Joyce, James M. 1998. "A Nonpragmatic Vindication of Probabilism." *Philosophy of Science* 65:575–603.

—. 1999. *The Foundations of Causal Decision Theory*. Cambridge, UK: Cambridge.

—. 2009. "Accuracy and Coherence: Prospects for an Alethic Epistemology of Partial Belief." In Franz Huber and Christoph Schmidt-Petri (eds.), *Degrees of Belief*, 263–297. Springer.

Konek, Jason and Levinstein, Ben. forthcoming. "The Foundations of Epistemic Decision Theory." *Mind* .

Lasonen-Aarnio, Maria. 2015. "New Rational Reflection and Internalism about Rationality." In Tamar Szabó Gendler and John Hawthorne (eds.), *Oxford Studies in Epistemology, Volume 5*, 145–171. Oxford: Oxford.

Leitgeb, Hannes and Pettigrew, Richard. 2010. "An Objective Justification of Bayesianism II: The Consequences of Minimizing Inaccuracy." *Philosophy of Science* 77:236–272.

Lewis, David. 1971. "Immodest Inductive Methods." *Philosophy of Science* 38:54–63.

—. 1980. "A Subjectivist's Guide to Objective Chance." In R. C. Jeffrey (ed.), *Studies in Inductive Logic and Probability*, volume 2, 263–93. Berkeley: University of California Press.

Mayo-Wilson, Conor and Wheeler, Gregory. 2016. "Scoring Imprecise Credences: A Mildly Immodest Proposal." *Philosophy and Phenomenological Research* 93:55–78.

Meacham, Christopher J. G. and Weisberg, Jonathan. 2011. "Representation Theorems and the Foundations of Decision Theory." *Australasian Journal of Philosophy* 89:641–663.

Pettigrew, Richard. 2016. *Accuracy and the Laws of Credence*. Oxford, UK: Oxford.

Railton, Peter. 1984. "Alienation, Consequentialism, and the Demands of Morality." *Philosophy & Public Affairs* 13:134–171.

Savage, Leonard. 1972. *The Foundations of Statistics*. New York: Dover. Second edition.

Schoenfield, Miriam. 2015. "Bridging Rationality and Accuracy." *Journal of Philosophy* 112:633–657.

Titelbaum, Michael G. 2015. "Rationality's Fixed Point (Or: In Defense of Right Reason)." In Tamar Gendler and John Hawthorne (eds.), *Oxford Studies in Epistemology, Volume 5*, 253–293. Oxford: Oxford.

Wedgwood, Ralph. 2013. "Rational 'Ought' Implies 'Can'." *Philosophical Issues* 23:70–92.

Williams, Bernard. 1981. "Internal and External Reasons." In *Moral Luck*, 101–113. Cambridge: Cambridge.