

The Rejection of Epistemic Consequentialism

Selim Berker
Harvard University
sberker@fas.harvard.edu

[Penultimate draft of a paper that was eventually published in
Philosophical Issues 23 (2013): 363–87; please cite that version.]

I. Introduction

“To later generations,” Christine Korsgaard tells us, “much of the moral philosophy of the twentieth century will look like a struggle to *escape* utilitarianism. We seem to succeed in disproving one utilitarian doctrine, only to find ourselves caught in the grip of another” (Korsgaard 1993, 275). Philippa Foot agrees: “It is remarkable how utilitarianism tends to haunt even those of us who will not believe in it. It is as if we for ever feel that it must be right, even though we insist that it is wrong” (Foot 1985, 196). I believe that a similar specter haunts contemporary epistemology: the specter of reliabilism. This essay represents my attempt at an exorcism.¹

A more accurate name for reliabilism is “truth-conducivism.” The guiding thought behind reliabilism² is that what makes a belief epistemically justified is its connection to a process (or indicator, or method, or . . .) that conduces toward the promotion of true belief and the avoidance of false belief. This thought is so alluring that many philosophers who do not consider themselves to be reliabilists nonetheless hold that truth-conduciveness is the defining mark of epistemic justification. Witness, for example, the common complaint that coherentism is untenable unless it can be shown that coherent beliefs tend to be mostly true. Or witness the speed with which discussions of the trustworthiness of philosophical intuitions about cases morph into discussions of the truth-conduciveness of such intuitions. Or witness the fact that many ardent anti-reliabilists allow that, although a belief’s being formed via a reliable process does not suffice for *prima facie* justification, a lack of evidence that one’s belief was formed via an *unreliable* process

¹ Except for one major difference near the end (see n. 22), the argument to follow overlaps a good deal with the main argument from my “Epistemic Teleology and the Separateness of Propositions.” I have tried, as much as possible, to avoid plagiarizing my past self, but some turns of phrase and even entire sentences from that previous paper have managed to work their way into this one.

² Unless otherwise specified, by “reliabilism” I mean “reliabilism about epistemic justification.” Most of what I say in this paper concerning reliabilism about epistemic justification also applies, with minor modifications, to reliabilism about knowledge, reliabilism about epistemic virtue, reliabilism about rational belief, and so on.

is necessary for all-things-considered justification. To echo Korsgaard's language, even as we forswear one reliabilist strain of thought, we find ourselves caught in the grip of another.

For Korsgaard and Foot, the inexorable pull of utilitarianism stems from its being a species of a certain genus, namely consequentialism. The same is true, I believe, of reliabilism. When I say this, I don't just mean that reliabilism, like utilitarianism, is a species of a certain genus which is the true source of that species' appeal (although I do mean that as well). No, what I mean is that reliabilism is in fact a species of the *same* genus as utilitarianism: reliabilism is also a version of consequentialism. So really we have *one* specter haunting *both* ethics and epistemology.

All consequentialist theories hold that (i) there are certain states of affairs that are inherently good, and other states of affairs that are inherently bad, and (ii) all other normative notions under the theory's purview are determined by how well the objects of assessment conduce toward or promote those states of affairs (either by directly conducing toward them, or by indirectly conducing toward them, in virtue of being suitably connected to some other entity that directly conduces toward them). If we apply this general idea to the ethical realm and take the value to-be-promoted to be individual well-being (with no regard to how that well-being is distributed), what results is utilitarianism. If we apply this general idea to the epistemic realm and take the value to-be-promoted to be true belief and the avoidance of false belief, what results is reliabilism. What matters epistemically, for reliabilists, are the *consequences* (in terms of attaining true beliefs and avoiding false beliefs) of forming beliefs in a certain way.

The idea that reliabilism is a form of epistemic consequentialism is not new. In his reliabilist manifesto, "What Is Justified Belief?," Alvin Goldman explicitly took the act-utilitarian analysis of RIGHT ACTION to be the model for his process-reliabilist analysis of JUSTIFIED BELIEF (Goldman 1979, 105). And in his presidential address delivered at the Annual Eastern Meeting of the American Philosophical Association a year after the publication of Goldman's article, Roderick Firth characterized reliabilism as a form of "epistemological rule-utilitarianism" (Firth 1981, 264, emphasis omitted). In §II, below, I say more about why I think Firth was exactly right in this characterization.

But I think that Firth was right about much more than that. In that presidential address, and in a

series of lectures delivered at the University of Arizona two years earlier (Firth 1978), Firth unveiled a way of arguing against certain brands of epistemic consequentialism that has, I think, been underappreciated. This lack of appreciation is, I suspect, partially due to historical happenstance. But it is also due, in part, to the fact that Firth's case against epistemic consequentialism was incomplete. What he provided us was really only the opening move in a more complicated argument. In §III, below, I show how we can fill in the details of Firth's argument. What emerges is not just a way of arguing against truth-conducivism, in its many manifestations, but more generally a way of taking on the entire epistemic-consequentialist framework that makes reliabilist strains of thought seem so inevitable.

II. Consequentialism, Whether Ethical or Epistemic

Traditionally, consequentialist theories are held to consist of two parts: an "axiology," which evaluates the goodness and badness of certain states of affairs, and a "deontology," which assigns deontic properties such as *being permissible* and *being obligatory* on the basis of the theory's axiology. Neither piece of terminology is ideal. For one thing, "deontology" has, rather confusingly, also come to be the name for what is supposed to be the chief sort of view opposing consequentialism.³ For another thing, I think it is important to distinguish between two different ways in which consequentialists theorize about the extension of evaluative properties such as *being good* and *being bad*, and the term "axiology" suggests a unity that irons over this difference.

To avoid such confusion, I prefer to break consequentialist theories into three components that I call "a theory of final value," "a theory of overall value," and "a deontic theory" (my replacement for "deontology"). The first of these specifies certain states of affairs that have value as ends in themselves, and other states of affairs that have *dis*value as ends in themselves. Note that I say here "have value as ends in themselves" and not "have value in virtue of their intrinsic properties": as Korsgaard and others have taught us (Korsgaard 1983, Rabinowicz and Rønnow-Rasmussen 1999), the final vs. instrumental value distinction should not be conflated with the intrinsic vs. extrinsic value distinction, and it is the

³ Or at least it has in ethics. In epistemology, "deontology" has, as a result of William Alston's unfortunate choice of terminology (Alston 1985, 1988), come to refer to a position that is orthogonal to the epistemic consequentialism vs. epistemic non-consequentialism distinction. (For more discussion, see my "Epistemic Teleology and the Separateness of Propositions," pp. 42–43.) So really we have three different uses of the term "deontology" in contemporary philosophy.

former that is most relevant in the current context. As long as the values in question are final values, and as long as the bearers of those values are states of affairs, I mean to be maximally permissive as to the contents of what I am calling “a theory of final value.” Some theories of final value are monist, positing only one sort of state of affair that has value as an end in itself; others are more pluralist. Some theories of final value assign degrees of value; others are less precise. Some theories of final value appeal to a distinction between *pro tanto* and all-things-considered value; others do not. Some theories of final value countenance agent- and time-relative values; others deal only with agent-neutral, time-neutral values.⁴ And so on.

The second component of any consequentialist theory is “a theory of overall value,” which assigns a value ranking to any entity that conduces toward or promotes those states of affairs which, according to the consequentialist’s theory of final value, have value or disvalue as ends in themselves. More precisely, since a consequentialist may want to evaluatively rank many different sorts of entities (acts, motives, rules, institutions, etc.), her theory of overall value might break up into several different sub-theories corresponding to each of these “evaluative focal points,” as Shelly Kagan (2000) calls them. For each evaluative focal point, the theory of overall value specifies a comparative ranking of the items falling within that evaluative focal point in terms of how well they conduce toward or promote the states of affairs featured in the consequentialist’s theory of final value. I mean to be maximally permissive as to what counts as a way of conducting toward or promoting a state of affairs. Most theories of overall value take the conducting relation to encompass *casual means*. But some theories of overall value also allow non-causal means of promotion, such as *constitutive means* or (if the constitution relation is distinct from the metaphysical grounding relation) *grounding means*.⁵ And there are theories of overall value that restrict the conducting relation in various ways, so that, for example, only *proximate* causal means are relevant, or only causal means in which the effect is an *intended* (and not merely foreseen) *consequence*. I also mean to be

⁴ There are authors who build into their definition of consequentialism a commitment to agent-neutrality, but I think this convention should be resisted: some agent-relative theories share the basic structural features possessed by agent-neutral consequentialist theories, whereas other agent-relative theories do not, and it is useful to represent this difference by allowing the former sort of theory to qualify as a form of consequentialism.

⁵ State of affairs X is an (upwards) constitutive means to state of affairs Y iff X’s obtaining (at least partially) constitutes Y’s obtaining. State of affairs X is a grounding means to state of affairs Y iff X’s obtaining (at least partially) makes it the case that Y obtains (or, equivalently, iff Y obtains [at least partially] in virtue of X’s obtaining).

maximally permissive as to the nature of the evaluative ranking that results for each evaluative focal point: it can be cardinal or ordinal, total or partial, etc.

The final component of a consequentialist theory is “a deontic theory,” which assigns deontic properties such as *being right/wrong*, *being justified/unjustified*, and *being permitted/forbidden* on the basis of the consequentialist’s theory of overall value.⁶ As before, since a consequentialist may want to deontically assess many different sorts of entities (motives as well as acts, rules as well as institutions, and so on), her deontic theory might break up into several different sub-theories corresponding to each of these “deontic focal points.” Some deontic sub-theories *directly* assign deontic properties on the basis of overall evaluative properties, as in act-consequentialism. But others take the form of an *indirect theory*—like what we find in rule-consequentialism, in which deontic properties are *directly* assigned to rules and then *indirectly* to acts on the basis of a bridge principle linking the deontic properties of rules to the deontic properties of acts. Finally, no assumption is made here that a deontic sub-theory is *maximizing* in form: it could be a *satisficing* theory, or something more complicated than a pure maximizing or satisficing theory.

Perhaps the best way to illustrate my tripartite way of characterizing consequentialist theories is through a few examples. First, two familiar examples from the ethical realm:

maximizing hedonistic act-utilitarianism:

i. *theory of final value:*

Pleasurable experiences⁷ have final value.
Painful experiences have final disvalue.
Nothing else has final value or disvalue.

ii. *theory of overall value* (for evaluative focal point {acts}):

S’s ϕ -ing at time t has more overall value than S’s ψ -ing at t iff the net balance of pleasure over pain that would be brought about if S ϕ -s at t is greater than the net balance of pleasure over pain that would be brought about if S ψ -s at t .

iii. *deontic theory* (for deontic focal point {acts} and deontic property *being right*):

S’s ϕ -ing at time t is right iff no other act available to S at t has more overall value.

⁶ I assume here, contra Alston (1985, 1888), that *being epistemically justified* is a deontic property. If you agree with Alston, simply replace all occurrences of “deontic” in this paper with “deontic and/or justificatory.”

⁷ Here “pleasurable experience” should be understood as shorthand for “a state of affairs in which someone has a pleasurable experience.” Similarly for talk of “painful experiences,” “true beliefs,” and “false beliefs” in what follows.

satisficing hedonistic rule-utilitarianism:

i. *theory of final value:*

Pleasurable experiences have final value.
Painful experiences have final disvalue.
Nothing else has final value or disvalue.

ii. *theory of overall value* (for evaluative focal point {rules}):

Set of rules r has more overall value than set of rules r' iff r tends (when internalized by a suitable portion of the population in a suitable range of circumstances) to lead to a greater balance of pleasure over pain in the world than r' does.

iii. *deontic theory* (for deontic focal point {rules} and deontic property *being right*):

Set of rules r is right iff r 's overall value is greater than or equal to some threshold, T.

deontic theory (for deontic focal point {acts} and deontic property *being right*):

S's ϕ -ing at time t is right iff the internalized set of rules that caused S to ϕ at t is right.⁸

The application of this framework to the epistemic realm involves one crucial change: rather than dealing with value *simpliciter* in their theories of final and overall value, epistemic consequentialists usually postulate that there is a special type of value, so-called “epistemic value,” which their consequentialist apparatus is concerned with the promotion of. With that in mind, here is a way of fitting a simplified⁹ version of process reliabilism into my tripartite framework:

process reliabilism:

i. *theory of final value:*

True beliefs have final epistemic value.¹⁰
False beliefs have final epistemic disvalue.
Nothing else has final epistemic value or disvalue.

ii. *theory of overall value* (for evaluative focal point {belief-forming processes}):

Belief-forming process b has more overall epistemic value than belief-forming process b' iff b tends (when employed in a suitable range of circumstances) to yield a greater ratio of true to false beliefs than b' does.

⁸ Some rule-consequentialists appeal to a different sort of bridge principle between the deontic properties of rules and the deontic properties of acts than the one featured here. For these rule-consequentialists, an act is right iff it *conforms to* a right set of rules (regardless of whether the agent has internalized those rules), whereas according to the bridge principle featured here, an act is right iff it is *done on the basis of* a right set of rules. I have opted for the latter to make the analogy with process reliabilism tighter.

⁹ “Simplified” because it neglects the distinction between belief-dependent and belief-independent cognitive processes. (See Goldman 1979, 116–17.) Nothing I go on to say turns on this distinction, so I have omitted it for simplicity’s sake.

¹⁰ Note that I say here “final epistemic value” and not “epistemic final value.” This is to allow for the possibility that although true beliefs have final value with regard to the epistemic domain, they might only have that value in virtue of conducing toward some other value outside of the epistemic domain, as in Kornblith 1993 and Sosa 2007, lect. 4.

- iii. *deontic theory* (for deontic focal point {belief-forming processes} and deontic property *being reliable*):

Belief-forming process *b* is reliable iff *b*'s overall epistemic value is greater than or equal to some threshold, *T*.

deontic theory (for deontic focal point {beliefs} and deontic property *being justified*):

S's belief that *p* at time *t* is justified iff the belief-forming process that caused S to believe that *p* at *t* is reliable.

The striking similarities between *satisficing hedonistic rule-utilitarianism* and *process reliabilism* should be enough to vindicate Firth's claim that reliabilism is a form of "epistemological rule-utilitarianism." The former is, in effect, pleasure-over-pain-conducivism with regard to action-guiding rules (and, by proxy, the actions so guided), the latter truth-over-falsity-conducivism with regard to belief-forming processes (and, by proxy, the beliefs so formed).

But many more epistemic theories other than process reliabilism qualify as versions of epistemic consequentialism. Alvin Plantinga's proper functionalism (Plantinga 1988, 1993), Ernest Sosa's virtue epistemology from his *Knowledge in Perspective* days (Sosa 1991), Richard Foley's theory of epistemic rationality (Foley 1987, 1993), William Alston's epistemic desiderata approach (Alston 1993a, 2005)—all can, with minimal pinching, be shoehorned into my three-component framework. Indeed, although I lack the space to fully defend this claim here,¹¹ there is a way of viewing much of the debate between externalists and internalists about epistemic justification over the past few decades as a debate *within* epistemic consequentialism, roughly corresponding to the debate between objective and subjective consequentialists in the ethical realm. The idea that epistemic normativity is fundamentally a matter of what conduces toward the promotion of epistemic value is pervasive among epistemologists of our day. And it is this pervasiveness which explains, I believe, why so many who renounce reliabilism nonetheless find themselves sliding into defending theses that are essentially reliabilist in spirit, if not letter.

As a result, I believe the best strategy for resisting this tendency is not to argue against the details of individual reliabilist proposals, but rather to take on the entire epistemic-consequentialist framework of which reliabilism is only the most prominent exemplar. Once we see that this framework is fundamentally misguided—once we see that it is not the case that our most basic epistemic commitment is to bringing

¹¹ I go into more detail in §III of my "Epistemic Teleology and the Separateness of Propositions."

about a world that contains more states of affairs with a distinctively epistemic sort of value—reliabilist strains of thought will no longer seem so inevitable. The problem with truth-conducivism is that it is a form of epistemic-value-conducivism, and conducivism of any sort is the wrong way to think about epistemic normativity.

III. Firth's Case against Epistemic Consequentialism

John Doe is a brilliant set theorist who is on the cusp of proving the Continuum Hypothesis: all he needs is six more months. But, alas, poor John is suffering from a serious illness that, according to his doctors, will almost certainly kill him in two months' time. John stubbornly clings to a belief that he will recover from his illness, and not only does this belief comfort him, but—let us suppose—it in fact significantly raises the chances that he will live for the six months that he needs both to complete his proof and to derive from it a variety of consequences for the rest of set theory. In other words, John's belief that he will recover is a causal means to his procuring a large number of true set-theoretic beliefs sometime in the future. But is John's belief epistemically justified? Is it the kind of belief that, from a purely epistemic perspective, he should be holding?

The story of John Doe comes from Roderick Firth's presidential address at the Eastern APA.¹² Firth expected his audience to agree with him that John's belief is not epistemically justified—that this is not the sort of thing that John should believe for purely epistemic reasons. The lesson Firth drew from this example is that, in his language, “instrumental epistemic merit” must be distinguished from “intrinsic epistemic merit.” Firth inferred from this that epistemic consequentialism—or “epistemological utilitarianism,” as he called it (Firth 1981, 262)—is untenable, since it in effect attempts to analyze intrinsic epistemic merit in terms of instrumental epistemic merit.

Firth's conclusion is essentially correct, I believe, but his inference was too quick. Firth's example of John Doe, and the others like it that he offered,¹³ are only counterexamples to extremely crude forms of epistemic consequentialism. Nevertheless, these examples point the way to a more serious problem that

¹² See Firth 1981, 259–60, as well as Firth 1978, 322–23. I have embellished Firth's original story in several ways.

¹³ See Firth 1978, 261, and 1981, 326 and 329.

afflicts almost all varieties of consequentialism with respect to epistemic matters.

Suppose our epistemic consequentialist accepts the theory of final value dubbed “veritism” by Alvin Goldman: namely, the view that (i) true beliefs, and only true beliefs, have final epistemic value, and (ii) false beliefs, and only false beliefs, have final epistemic disvalue (Goldman 1999, 5; 2001, 31). Suppose, moreover, that our epistemic consequentialist endorses a theory of overall value which allows all causal means to be ways of promoting final epistemic value. Then regardless of the other details of the epistemic consequentialist’s theory of overall value, and regardless of her deontic theory, it is difficult to see how our epistemic consequentialist can resist the conclusion that John Doe’s belief that he will recover is epistemically justified. This is most obvious if our epistemic consequentialist holds a *direct* deontic theory: because the epistemic pay-off, in terms of future true beliefs, of John’s belief is so great, that belief will have such a high overall epistemic value that, given almost any plausible way of directly assigning deontic properties to beliefs on the basis of their overall evaluative properties, it will follow that John is epistemically justified in believing that he will recover. However, if we tweak our story a bit, the same result holds even if our epistemic consequentialist embraces an *indirect* deontic theory: all we need to do is make situations like the one John faces sufficiently common in his world that a general practice of stubbornly believing that one will recover from an illness tends, on average, to greatly promote the acquisition of true beliefs.¹⁴

Many other forms of epistemic consequentialism face the same result. For example, suppose our epistemic consequentialist holds as her theory of final value what we might call “restricted veritism,” according to which the bearers of final epistemic value are not all true beliefs as a whole, but rather a certain class of true beliefs, namely the “important” or “interesting” or “significant” ones.¹⁵ This move is

¹⁴ Alternatively, we could increase the number of true beliefs that John Doe will acquire if he lives for six months to such a degree that, even if his situation is rare, his belief’s immense epistemic pay-off more than compensates for his situation’s rarity in the determination of whether, in general, a practice of forming beliefs-like-that in situations-like-that tends to promote final epistemic value.

¹⁵ Should an advocate of restricted veritism also restrict the class of false beliefs that have final epistemic disvalue to the important/interesting/significant ones, or can she allow that every false belief, regardless of its content, is epistemically bad as an end in itself? I leave this question open, since it won’t matter for my argument. But it is worth noting that the usual arguments for restricted veritism (we shouldn’t waste our epistemic lives counting grains of sand, etc.) only push us to restrict the range of true beliefs that have positive value, not the range of false beliefs that have negative value. (The possibility of only restricting one of these without restricting the other causes problems for a well-known argument against restricted veritism originally due to Marian David [2001, 159] and developed at greater length by Stephen Grimm [2009]. Another problem with David and Grimm’s

no help with Firth's John Doe case, since John's belief that he will recover and the set-theoretic beliefs he will thereby acquire are all important and interesting and significant. Similarly, suppose our epistemic consequentialist holds, as part of her theory of final value, that the disvalue of false beliefs is much greater—maybe even infinitely greater—than the value of true beliefs. Again, this move is no help with Firth-style cases, since (a) in the John Doe case, we can stipulate that not only will John form a large number of new true set-theoretic beliefs if he lives for six more months, but moreover he will revise a large number of previously held false set-theoretic beliefs, and (b) it is easy enough to construct variants of the John Doe example with a reverse structure, so that a belief which most of us would say is epistemically justified leads to a vast quantity of false beliefs, thus forcing the epistemic consequentialist who embraces veritism and allows causal means of epistemic-value promotion to say that this belief is in fact unjustified.¹⁶ Finally, suppose our epistemic consequentialist rejects veritism and takes cognitive states other than—or in addition to—true belief to have final epistemic value. Once again, this does not help with Firth-style cases. After all, not only will John Doe gain a plethora of true beliefs if he lives for another six months, but he will also acquire many new pieces of knowledge, much greater coherence in his web of beliefs, significantly more understanding, and maybe even some wisdom.

There are forms of epistemic consequentialism, though, that emerge unscathed when we consider Firth's original battery of examples. The basic formula for generating these cases should be apparent: simply take a standard example from discussions of whether there are practical reasons for belief, and switch the case so that rather than the belief in question being a causal means of promoting the believer's practical ends, it instead is a causal means of promoting her epistemic ends sometime in the future. This suggests two obvious ways of avoiding any untoward consequences when it comes to such cases. The first is to restrict the conducting relation in one's theory of overall value so that causal means are not allowed as

argument is that they neglect the possibility of cross-propositional value-promotion.)

¹⁶ Here is one such case. Suppose I am trying to find the location of a lecture given by an incredibly persuasive speaker who, unbeknownst to me, will be espousing a variety of false views. I ask a reliable-looking passerby for the lecture's location, and she says to me, "It's in room 210." Am I justified in believing that the lecture is in room 210? Plausibly, I am. But let us fill in the story as follows: if I attend the lecture, I will be dazzled by the speaker's sophisticated reasoning and rhetorical flourishes, and will leave the lecture hall firmly believing a wide array of falsehoods. Given the dire epistemic consequences of attending this lecture, it seems that the epistemic consequentialist who accepts veritism and countenances causal means must say that I am *not* epistemically justified in believing that the lecture is located in room 210, since that belief (or the process that formed it) would be a causal means to my acquiring a large number of false beliefs.

a way of promoting final epistemic value; for example, maybe only constitutive means count as a way of conducing toward value of that sort. The second obvious way of defusing these cases is to time-relativize one's theory of theory of final value, so that—at least in the case of veritism—only states of affairs in which a subject *at that time* truly believes a proposition are deemed to have final epistemic value for a given subject at a given time, and similarly for the disvalue of false beliefs.¹⁷ Thus, it would appear, epistemic consequentialists have at least two ways to avoid Firth's criticisms.

It is here that we need to go beyond anything Firth explicitly argued. First, we need to think a bit harder about what, exactly, is doing the work in Firth-style cases. Why, for example, do many forms of epistemic consequentialism yield the verdict that John Doe's belief is epistemically justified? At first pass, it is because of a certain hallmark of consequentialist theories in general: namely, their proclivity for trade-offs. The values and disvalues that make up a consequentialist's theory of final value can, almost always, compete and conflict with another.¹⁸ Where such competitions and conflicts arise, the standard consequentialist way of settling the matter is through aggregation: the theory of overall value somehow “adds up” the positive final values and “subtracts” the negative final values promoted by each item in a given evaluative focal point to get a total that can be compared to the total for every other time in that focal point. This leads to the infamous tendency of consequentialist theories to countenance trade-offs when determining what one should do, believe, and so on. I say “infamous” because this penchant for allowing—or even mandating—trade-offs is what gives rise to many of the most notorious intuitive counterexamples to consequentialist theories in the practical realm: cases in which consequentialist theories seem to require us to do things such as cut up the one to save the five (Thomson 1976, 80), or kill someone to spare thousands from a minor headache (Norcross 1997, 135), or allow a technician to continue to be painfully shocked so that millions can watch the end of a soccer match (Scanlon 1998, 235). From the perspective of many forms of ethical consequentialism, each of these cases involves a trade-off in which a multitude of good benefits outweigh a single unfortunate outcome. Similarly, from the

¹⁷ This, in effect, is what Richard Foley urges us to do in the opening pages of his 1987 to handle Firth-style cases.

¹⁸ Indeed, if conflict and competition were not possible among the final values, there would be no need for a theory of overall value within the consequentialist framework: we could (in the case of act-consequentialism) simply define an action as obligatory if it brings about a good state of affairs, forbidden if it brings about a bad state of affairs, and so on.

perspective of many forms of epistemic consequentialism, each of Firth's cases involves a trade-off in which a multitude of good epistemic benefits outweigh a single epistemically unfortunate outcome.

But the parallel is deeper than that. In the practical domain, most of the problematic trade-off cases that have been pressed against consequentialism involve trade-offs that "cross the boundary between persons": they involve a decision to harm or hurt one person in order to benefit or advantage a number of other people; hence the charge that many forms of consequentialism "ignore the separateness of persons," by treating inter-personal trade-offs as fundamentally no different from intra-personal ones. It is noteworthy that, when we turn to the epistemic domain, all of Firth's cases involve trade-offs that, as we might put it, "cross the boundary between propositions": they involve a subject allowing an epistemically bad cognitive state with a certain propositional content to occur in order to promote a large number of epistemically good cognitive states with propositional contents distinct from that one.¹⁹ So maybe the true diagnosis of why many forms of epistemic consequentialism have problems with Firth-style cases is that they, as it were, "ignore the separateness of propositions": they take the epistemic status of a given belief to partially turn on whether it conduces (either directly or indirectly) toward the promotion of epistemically good and epistemically bad cognitive states with a different propositional content.

Let us work with this hypothesis for the time being: that the reason some epistemic consequentialists run afoul of Firth-style cases is that they ignore the separateness of propositions. As it turns out, I don't think this is the entire story, but it will take some work before I can explain why it is only partially correct. (What I am eventually going to argue is that, although it is indeed an error to ignore the separateness of propositions, this is just an instance of a more general error.) Right now, though, we are working with the hypothesis that the reason why many epistemic consequentialists are susceptible to Firth-style counterexamples is precisely because they countenance cross-propositional trade-offs in their determination of the epistemic status of individual beliefs. The crucial thing to notice is that, if this is

¹⁹ If an epistemic consequentialist takes some of the bearers of final epistemic value to be cognitive states that do not have propositional content (such as objectual understanding, which some theorists deny is a relation between a subject and anything propositional), then the claim in this sentence needs to be slightly reformulated, but an analogous point will hold (for example, instead of talk of trade-offs that "cross the boundary between propositions," we might need to talk of trade-offs that "cross the boundary between domains of inquiry"). In what follows, I assume that all of the relevant cognitive states have propositional content, mainly for ease of exposition. Nothing turns on this.

correct, then even epistemic consequentialists who restrict the conducting relation in their theory of overall value to rule out all causal means and even epistemic consequentialists who embrace a time-relative theory of final value are in trouble, for they still allow for the possibility that cross-propositional trade-offs can impact the epistemic status of an individual belief. Epistemic consequentialists of either of these two sorts avoid having to say embarrassing things about any of the original cases bequeathed to us by Firth; but, if we are sufficiently crafty, we can find other cases which these sorts of epistemic consequentialists are forced to say embarrassing things about due to the presence of cross-propositional trade-offs.

The cleanest case of this sort that I know of involves making a certain potentially contentious assumption in the philosophy of mind. The assumption is this: that second-order beliefs about one's own beliefs are partially constituted by the first-order beliefs they are about, so that the relationship a given first-order belief stands to its corresponding second-order belief is that of a constitutive means.²⁰ I call this assumption "potentially contentious" since some theorists hold that, instead, a first-order belief and its second-order cousin are distinct existences and thus, at most, the former is a causal means to the latter. For the purposes of the example to follow, I will just be assuming that these theorists are wrong. (In the end, this assumption will be dispensable.)

On to the example. Suppose that I happen to be an extremely self-conscious creature, so that whenever I form a belief about the external world, I also—at the same time—form a variety of second-order beliefs about this belief and the way in which it was formed, a variety of third-order beliefs about each of those second-order beliefs and the ways in which they were formed, a variety of fourth-order beliefs about each of those third-order beliefs and the ways in which they were formed, and so on. (To appease those who think that finite creatures can't have an infinite number of beliefs, let us say that this hierarchy stops at the sixth order.) Moreover, although I am excellent at introspection, and almost always have accurate higher-order beliefs about the contents of my lower-order beliefs and how I came to hold them, let us stipulate that I am just terrible at forming beliefs about the external world: I almost always get things wrong. So, for instance, one day I come to (falsely) believe *that the world will end in 2013* by consulting

²⁰ More precisely, the assumption is that second-order beliefs about one's own beliefs *formed at least partly via introspection* are partially constituted by the first-order beliefs they are about. This qualification should be taken as understood in what follows.

my trusty tea leaves. At the same time, I form the following second-order beliefs: “I believe *that the world will end in 2013*,” “I came to believe *that the world will end in 2013* by consulting tea leaves,” “I first came to believe *that the world will end in 2013* on a Monday,” and so on. I also, at the same time, form the following third-order beliefs: “I believe *that I believe that the world will end in 2013*,” “I came to believe *that I believe that the world will end in 2013* through introspection,” “I first came to believe *that I believe that the world will end in 2013* on a Monday,” and so on. This continues until the sixth level.

I take it to be obvious that my first-order belief in the world’s imminent demise is unjustified. (Fans of tea-leaf prognostication can swap in their favorite example of an ill-conceived way of forming beliefs about the external world.) The problem for epistemic consequentialists is that this belief was a constitutive means of, at the same time, acquiring an immense number of true beliefs (and, for that matter, vast amounts of new knowledge, greatly increased coherence in my web of beliefs, much deeper understanding of my belief-forming habits, and so on). So even if our epistemic consequentialist endorses a time-relative conception of final epistemic value, and even if she restricts the conducing relation in her theory of overall value so that it only includes constitutive means, she is forced to say that my first-order belief—or (in the case of an indirect deontic theory) the process/practice/method by which it was formed—involves a sacrifice for the greater epistemic good, and hence the belief is justified.²¹

This case shows that the two obvious responses we have been considering to Firth’s original style of counterexample only put off the inevitable: even epistemic consequentialists who time-relativize their final values and/or restrict themselves to causal means ignore the separateness of propositions by allowing the epistemic status of a given belief with a given propositional content to turn on the degree to which it promotes (either directly or indirectly) epistemically good or bad cognitive states with a different propositional content. Moreover, once we notice the structure of the example I have provided, we can see that the assumption I was making in the philosophy of mind was not essential to it. As long as the epistemic consequentialist allows constitutive means to be a way of promoting epistemic value, it should

²¹ It is also possible to construct cases like this with a reverse structure, in which a belief that most of us would deem to be justified serves as a constitutive means to, at the same time, forming a host of false higher-order beliefs (and other epistemically bad higher-order cognitive states).

be possible to construct problem cases of this sort by appending simultaneously-occurring constitutive means upon simultaneously-occurring constitutive means in a tree-like manner.

This same dialectic replays itself, I believe, with nearly every attempt to complicate one's brand of epistemic consequentialism in order to avoid Firth-style cases. For example, consider the case of process reliabilism. When process reliabilists talk of the truth-ratio of the beliefs yielded by a given process, they mean to speak of the beliefs which that process *directly* yields, without any causal intermediaries other than the belief-forming process itself. Thus, in effect, process reliabilists restrict the conducting relation in their theory of overall value to *proximate causal means*. This feature of their view would appear to allow process reliabilists to avoid any troubling consequences when it comes to either Firth's John Doe case (since they deny that non-proximate causal means promote final epistemic value) or my incredibly accurate introspector case (since they deny that constitutive means promote final epistemic value). But, as before, this just puts off the inevitable, for process reliabilists still allow cross-propositional trade-offs to affect the epistemic status of a given belief, as long as these trade-offs occur between that belief and the other beliefs (with differing propositions contents) that can be formed via the same process. So it is to such trade-offs that we must turn in order to construct a problem case for process reliabilists.

Here is one such case. Suppose the following is true of me: whenever I contemplate whether a given natural number is prime, I form a belief that it is not. "Is 25 prime? No, it is not." "Is 604 prime? No, it is not." "Is 7 prime? No, it is not." Let us also stipulate that this is the only cognitive process by which I form beliefs about the primeness of natural numbers. (I'm a simpleminded kind of guy.) Since the ratio of prime to composite numbers less than n approaches 0 as n approaches infinity, my belief-forming process tends to yield a ratio of true to false beliefs that approaches 1. Therefore process reliabilists are forced to say that, because my belief-forming process is almost perfectly reliable, any belief formed on its basis is justified. But that's crazy! When I form a belief *that 7 is not prime*, it is simply not correct to say that, although that belief is false, it is epistemically redeemed by the truth of the other beliefs which would be formed via the process that led to it. Thus even process reliabilists countenance cross-propositional trade-offs, and where they do, it seems their theory yields the wrong epistemic verdict.

Let us consider one last consequentialist gambit.²² All of the cases—whether mine or Firth’s—that I have used so far to demonstrate the inadequacies of various incarnations of epistemic consequentialism share a common feature: they involve trade-offs which cross the boundary between propositions, as I have been putting it. So maybe what epistemic consequentialists need to do is build into their theory, by fiat, adherence to the separateness of propositions. One way of doing this would be for epistemic consequentialists to add two restrictions to their theory of overall value.²³ First, they should restrict their theory of overall value to evaluative focal points in which each item in that focal point has a specific propositional content; in the case of the evaluative focal point {beliefs}, this restriction yields no change, but in the case of the evaluative focal point {belief-forming processes}, it requires us to consider belief-forming processes which are narrowly enough individuated that something only counts as an instance of that process if it leads to a belief with one specific propositional content (for example: the process *forming a belief that there is a computer in front of one via visual perception*, rather than the process *forming a belief about the external world via visual perception*). Second, epistemic consequentialists should restrict the conducting relation in their theory of overall value so that an item in an evaluative focal point with a given propositional content only counts as conducting toward an epistemically valuable end if that epistemically valuable end *has the same propositional content*. Making these two restrictions will bar cross-propositional trade-offs from ever being relevant to the epistemic standing of an individual belief.

It is worth noting an immediate cost of embracing these two restrictions. Epistemic consequentialists who do so while accepting a veritistic theory of final value will have a difficult time averting the consequence that every necessarily-true belief, regardless of the basis on which it is held, counts as epistemically justified.²⁴ (The class of necessarily-true beliefs includes both beliefs in necessarily-true propositions, such as a belief *that anything which is actual is also possible*, and also self-confirming beliefs,

²² The material to follow, and the resulting diagnosis (which goes beyond accusing epistemic consequentialists of ignoring the separateness of propositions), represents a departure from my “Epistemic Teleology and the Separateness of Propositions.”

²³ It might be thought that another way of imposing the separateness of propositions on a consequentialist epistemic theory is to move to a proposition-relative theory of final value. But, as I explain in my “Epistemic Teleology and the Separateness of Propositions,” pp. 37–38, such a move, with no accompanying changes in one’s theory of overall value, still allows cross-propositional trade-offs to play a role in the determination of an individual belief’s epistemic status.

²⁴ I suspect that epistemic consequentialists who hold a non-veritistic theory of final value will face analogous problems when it comes to the epistemic evaluation of necessarily-true beliefs, though I shan’t argue for that here.

such as a belief *that there are beliefs*.) This consequence is clearest in the case of process reliabilism. Within the confines of process reliabilism, the maneuver we are considered is equivalent to adding to process reliabilism a constraint that each belief-forming process only yields beliefs with a single propositional content. And it is well known that when process reliabilists individuate belief-forming processes in this narrow a manner, they are forced to deem even the most haphazardly-formed necessarily-true belief as being epistemically justified (Conee and Feldman 1998, 149). Moreover, a similar result clearly holds when we consider a consequentialist epistemic theory that (a) includes a version of veritism as its theory of final value, (b) restricts its theory of overall value in the two ways needed to avert cross-propositional trade-offs, and (c) features a direct deontic theory (rather than an indirect one, as in process reliabilism).

However, it is not just necessarily-true beliefs that epistemic consequentialists who make these two restrictions have trouble with. It turns out, in fact, that they even have trouble with the belief featured in Firth's original John Doe case—although why they do is not immediately obvious.

It will help to focus on a slightly different case for a moment. So let us consider John's cousin, Jane Doe, who is a brilliant chess master six months away from playing the most important tournament of her life. But, alas, poor Jane has been diagnosed with the same illness as John (apparently it is congenital), and her doctors also give her only two months to live. Like John, Jane stubbornly clings to a belief that she will recover from her illness, and, as in John's case, let us suppose that this belief in fact significantly raises the chances that she will live long enough to play in her tournament six months from now.

Examples involving figures like Jane Doe pre-date Firth's John Doe variant of them by a number of decades. Usually a scenario such as Jane's is brought up in order to ask: even though Jane's belief that she will recover is not justified on epistemic grounds, does she nonetheless have good reason to hold that belief in virtue of the practical benefits of doing so? That is, can there be practical reasons for belief? Firth's innovation was to take this familiar sort of example and to switch the ends that are furthered by the subject's belief from being practical ends (as in the case of Jane the chess master) to being epistemic ends (as in the case of John the set theorist). But it is interesting to note that, even in the earlier Jane Doe case, there are *some* epistemic ends that are at stake. For, by stipulation, Jane's belief that she will recover raises

the chances that she will in fact recover. Or in other words: Jane's belief that she will recover raises the chances that she will thereby have a true belief. Or in other words: Jane's belief that she will recover promotes her being in a cognitive state which (according to advocates of veritism) has final epistemic value, where this cognitive state has the same propositional content as the state that promotes it. But, I insist, Jane's belief is not for that reason epistemically justified.

This last claim is most obvious when we consider a version of the Jane case in which she is not aware of the fact that her belief increases the likelihood of its own truth.²⁵ So let us suppose that she is not aware of this, and, to fix on some numbers, let us suppose that she has a 10% of recovering if she does not believe she will recover and a 90% of recovering if she does believe she will recover (where, moreover, the relevant percentages are *caused* by her being in the relevant doxastic state, not merely *correlated* with her so being). Then it appears that an epistemic consequentialist who accepts veritism is forced to conclude that Jane's belief is epistemically justified, because it conduces, either directly or indirectly, toward a state of affairs with final epistemic value, without promoting any compensating losses in epistemic value.²⁶ Moreover, and this is the crucial point, restricting our epistemic consequentialist's theory of overall value in the two ways suggested above does not alter this verdict about the case, since we are not dealing with cross-propositional value-promotion. So even cases not involving cross-propositional trade-offs cause trouble for epistemic consequentialists.

Now we are at the heart of the matter. Jane's belief makes it the case that she is more likely to have something which veritism deems to be epistemically good as an end in itself. But I take it to be obvious that her belief is not thereby epistemically justified. Thus the central problem with epistemic consequentialism is not that it neglects the separateness of propositions. Rather, the central problem is its focus on the promoting or conducing relation.

²⁵ I believe a parallel troubling result holds in a case in which she *is* aware of this fact, but arguing this takes a bit of care, since in at least some cases of this sort Jane has additional inductive reasons to hold her belief—and, by anyone's lights, these additional inductive reasons might be enough to make her belief justified. Nevertheless, even if Jane's belief that she will recover is justified in such a case, the epistemic consequentialist gives the wrong explanation of *why* her belief is justified: her belief is not justified because it increases the likelihood of its own truth, but rather because of the inductive reasons for its truth (which she could have even if the belief did not in fact increase the likelihood of its own truth this time around).

²⁶ As before, to secure this result in the case of an indirect deontic theory we may need to increase the frequency in Jane's world of cases in which forming a belief about one's health increases the likelihood that it will be true.

Consequentialism in ethics is famously forward-looking: it ties an action's, or rule's, or institution's ethical merit to the value of the states of affairs it helps bring about. Consequentialism in epistemology is also forward-looking (if not temporally, then at least in the order of explanation): it ties a belief's, or process's, or character trait's epistemic merit to the value of the states of affairs it helps bring about (whether causally or constitutively or otherwise). But this, I think, gets things exactly backwards. Consider the following slogans: "Epistemic justification is a matter of *responding to* how things appear to you," "Epistemic rationality is a matter of *respecting* one's evidence," "Epistemic virtue is a matter of *fitting together* one's cognitive states into a coherent whole." Each of these is, of course, highly metaphorical, and filling in the details of what these metaphors come to requires positive theorizing. But note that, in each case, epistemic normativity is characterized as fundamentally backward-, or at least sideways-, looking. Contrast that with epistemic consequentialism's central slogan: "Epistemic normativity is a matter of *bringing about* states of affairs with a distinctive sort of value." Here, and only here, do we have a forward-looking conception of what is mandated by epistemic norms. This explains, I believe, why epistemic consequentialists have such a difficult time making good on the earlier slogans.²⁷ It also explains, I believe, why so many forms of epistemic consequentialism ride roughshod over the separateness of propositions. And it explains, I believe, why even those versions of epistemic consequentialism that build into their theory a commitment to the separateness of propositions have trouble with the Jane Doe case—and, for that matter, with Firth's original John Doe case, since John Doe's belief that he will recover also serves to promote its own truth, and yet is not for that reason justified. It is epistemic consequentialism's fixation on the promotion of epistemically valuable state of affairs, its reduction of beliefs to mere instruments serving our independent epistemic ends, that causes all of these problem cases to arise. So Firth was right: epistemic consequentialism tries to analyze intrinsic epistemic merit in terms of instrumental epistemic merit, and for this reason should be rejected.

²⁷ I owe the observation that epistemic consequentialists have a difficult time accounting for the slogan "It is rational to respect your evidence" to Emily McWilliams.

IV. Concluding Remarks

I want to end by saying a few words about the nature and scope of my argument.

About the nature of my argument, I will be brief. In deploying my argument against consequentialist theories in epistemology, I have relied on a series of counterexamples—first an example due to Firth, and then a series of examples inspired by what I take to be the central insight behind Firth's example. However, I do not mean my argument to rest too heavily on these particular examples. The important point is not the examples themselves, but rather the *structure* of the examples. As I see it, my Firth-inspired argument against epistemic consequentialism involves not a series of counterexamples, but rather a *recipe* for generating counterexamples. For those consequentialist epistemic theories which violate the separateness of propositions, we can construct such a counterexample by finding a case in which the belief being epistemically evaluated in our deontic theory is sacrificed for the greater epistemic good, where that greater epistemic good involves cognitive states with a different propositional content from the one being epistemically evaluated. For those consequentialist epistemic theories which observe the separateness of propositions, we can construct a counterexample by finding a case in which the belief being epistemically evaluated in our deontic theory promotes, directly or indirectly, its own status as a cognitive state with final value. This recipe makes my style of argument particularly versatile: even if fault is found with the specific examples I have offered in this essay, my recipe allows me to find other examples of a structurally similar sort that suffice to make the same point.

For example, consider the prime number case I pressed against those process reliabilist theories which do not adhere to the separateness of propositions. A common response to this example is to insist that, in the case as I describe it, the relevant belief-forming process for my belief *that 7 is not prime* is not the same as the relevant belief-forming process for my belief *that 604 is not prime*, thus allowing the process reliabilist to avoid having to say that both the former and the latter are justified. Now, first of all, this reply strikes me as appealing to an all-too-convenient way of individuating belief-forming processes. But more importantly, this reply seems to concede that *if* my belief *that 7 is not prime* were formed via a process that has a high truth-ratio but only achieves that high truth-ratio by sacrificing any hope of attaining the

truth with regard to beliefs in the proposition *that 7 is not prime*, then my belief *that 7 is not prime* would not be justified. And this structural point is all I need in order to formulate my objection to those process reliabilists who do not impose the separateness of propositions on their theory, for surely *at least some* examples with this structure are possible when we individuate belief-forming processes wide enough that the same process may yield beliefs with distinct propositional contents.

What about the scope of my argument? In one way, its intended scope is very wide. Although I have not shown how to do so in every case, I believe it is possible to use my recipe to construct problem cases for all varieties of epistemic consequentialism that do not go so far as to restrict the conducting relation to the instantiation relation, or (even more radically) to give up on the epistemic evaluation of individual beliefs all together. But, in fact, the target of my argument is even larger than that. Before Elizabeth Anscombe introduced the term “consequentialism” in 1958 (Anscombe 1958, 12), consequentialism instead went by the name “teleology,” from the Greek *telos* for “aim” or “goal” (Broad 1930, 206; Muirhead 1932, 6). And although I have been characterizing the first component of a consequentialist theory as “a theory of the final value,” I just as easily could have characterized it as “a theory of ultimate goals” which specifies certain ultimate aims or fundamental goals that structure the norms under consideration: it is a fairly trivial matter to translate a theory of ultimate goals into a theory of final value, and vice versa. For this reason, I intend my argument also to apply to the host of epistemologists who take epistemic normativity to be ultimately explicable in terms of certain distinctively “epistemic” or “cognitive” or “intellectual” goals, such as the twin goals of acquiring true beliefs and avoiding false ones—provided, that is, that these goals are interpreted in a to-be-promoted/consequentialist manner.

So in some respects the intended target of my argument is quite broad; however, this last qualification reveals a way in which the target of my argument is actually fairly narrow. Strictly speaking, all I have argued is that epistemic normativity cannot be reduced to a set of epistemic values (or, equivalently, a set of epistemic goals) that are all “to be promoted.” Therefore it is compatible with my argument that epistemic values/goals play an (explanatorily) foundational role in epistemology, provided

that either (i) at least some of these epistemic values/goals are not of the “to be promoted” sort, or (ii) there is an additional foundational notion in epistemology not ultimately grounded in values/goals. I personally favor a more radical variety of epistemic non-consequentialism than either of these two options, namely one that does away with talk of “epistemic value” and “epistemic goals” all together. But I am not here today offering an argument against those who wish to hold onto such talk, so long as they either give it a non-consequentialist spin or append to their theory a non-consequentialist element in order to avoid the problematic sorts of cases discussed in this essay.

Thus it will vary from theorist to theorist, and from theory to theory, whether accepting everything I have argued here today entails a radical departure from business as usual. In some cases, a drastic overhaul is required. For example, a commitment to epistemic consequentialism runs so deep for reliabilists that I see no way of accepting my arguments and holding onto reliabilism. Similarly, William Alston is fond of saying that what distinguishes epistemic modes of normative appraisal from other modes of normative appraisal (moral, prudential, and so on) is that epistemic appraisal is “undertaken from what we might call the ‘epistemic point of view.’ That point of view is defined by the aim at maximizing truth and minimizing falsity in a large body of belief” (Alston 1985, 83). But I think my Firth-inspired argument definitively establishes that this way of characterizing epistemic modes of normative appraisal will not do: beliefs and belief-forming practices can do very well with regard to Alston’s aim without having the central sort of positive normative status we associate with epistemic norms. And this conclusion, of course, has ramifications for the rest of Alston’s work in epistemology (Alston 1989, 1991, 1993b, 2005), which often hinges on his teleological way of characterizing the epistemic point of view.

In other cases, embracing my argument brings with it a less radical shift in one’s views. For instance, Earl Conee and Richard Feldman’s brand of evidentialism is very much in spirit with the sort of epistemic non-consequentialism for which I have been arguing (Conee and Feldman 2004, 2008). Conee and Feldman do occasionally lapse into talk of “epistemic value” and “epistemic goals,” but I think this talk can be excised from their program without major loss. At one point, Feldman tentatively suggests that we can argue for evidentialism by noticing that it maximizes epistemic value to form one’s beliefs (and

other cognitive attitudes; hereafter I drop this qualification) in accordance with the dictates of evidentialism (Feldman 2000, 181–86). And at another point, Feldman proposes that what distinguishes epistemological duties from other sorts of duties is that the former subserve our epistemic goals, and thus conforming to these duties counts as a sort of epistemic success (Feldman 2002, 376–80). However, in the course of defending these two claims Feldman posits that our primary epistemic goal is rational or justified belief (Feldman 2000, 184–86; 2002, 379–80), and to handle Firth-style cases he, in effect, restricts the conducting relation to the instantiation relation (Conee and Feldman 2004, 257–58). But by the time he has done both of these, the consequentialist/teleological framework to which he is appealing has become toothless: it is doing no real argumentative work. Why are beliefs that fit one’s evidence epistemically justified? Because they maximize epistemic value. Why do they maximize epistemic value? Because only beliefs that fit one’s evidence are rational, and only rational beliefs have epistemic value. This argument will hardly persuade anyone not already won over to evidentialism. Why is a given duty to hold a given belief an epistemological duty, rather than a moral or prudential one? Because conforming to it subserves our epistemic goals. Why does conforming to it subserve our epistemic goals? Because that’s what you need to do in order to have justified beliefs, and having justified beliefs is our primary epistemic goal. Since “justified” here refers to *epistemic*, rather than *moral* or *prudential* justification, little explanatory progress has been made in this exchange. (We started out wondering what distinguishes epistemological from moral and prudential duties, and are left at the end wondering what distinguishes epistemic from moral and prudential justification.) Better, I think, for Conee and Feldman to jettison this talk of “epistemic value” and “epistemic goals” all together, since it is doing no real work in their theory.

Finally, there will be cases in which renouncing epistemic consequentialism does not have as extreme an impact as it does in the case of Alston’s project, or as little impact as it does when it comes to Conee and Feldman’s program: in these cases, shedding a tendency to think of epistemic normativity in terms of what conduces toward or promotes states of affairs with a distinctively epistemic sort of value will require reconfiguring the details of one’s theory without fundamentally shifting its central focus. For example, I think there is a way of reading the evolution of Ernest Sosa’s brand of virtue epistemology

according to which he has gradually shifted from a consequentialist conception of epistemic virtues in terms of truth-conduciveness (Sosa 1991, 2003) to a more truly Aristotelian way of conceiving of the epistemic virtues (Sosa 2007, 2011). And although much of the recent explosion of research on the value of knowledge and other cognitive states (Kvanvig 2003; Pritchard 2007; Haddock, Millar, and Pritchard 2009; Pritchard, Millar, and Haddock 2011) has, to my mind, implicitly assumed an essentially consequentialist way of thinking about value, I believe that many of the issues raised by this literature would survive, although in a slightly changed form, were it to be divested of its consequentialist overtones. These, however, are large issues, and I lack the space to properly develop these thoughts here. So let me end by mentioning an important challenge to the sort of epistemic non-consequentialism I am recommending in this essay.

The important challenge is this: what, exactly, is the connection between epistemic justification and truth, if that connection is not to be understood in terms of the former conducting toward the latter? The thought here is: “Surely there is some sort of an intimate link between a belief’s being epistemically justified and its being true.” (The first word in this thought is crucial: “surely” is philosopher-speak for “I have no argument for the claim I am about to make.”) Epistemic consequentialists who accept veritism have an account of this intimate link ready at hand: justification is connected to truth because justified beliefs conduce, either directly or indirectly, toward true beliefs. If we reject epistemic consequentialism, as I am urging that we do, we are left bereft of this seductively simple account of justification’s connection to truth. I concede that this is an important challenge for those of us who seek to pursue non-consequentialist options in epistemology.²⁸ In the end, I think there is no easy way to meet this challenge without relying on the details of one’s particular brand of non-consequentialist epistemic theory, and that is not a task I have undertaken in the current essay. But it is important to realize that, if the arguments of this essay succeed, then epistemic consequentialists who embrace veritism are no better off with regard to that challenge, since it follows from my arguments that their proposed explanation of the connection between justification and truth in terms of the conducting relation is inadequate. Consideration of the Jane Doe case helps bring this out. Epistemic consequentialists hope to explain the link between justification

²⁸ It is also an important challenge for those epistemic consequentialists who do not accept veritism.

and truth in terms of the very relation that Jane's belief that she will recover stands to that belief's truth. But since her belief is manifestly unjustified, this must be the wrong approach to take when attempting to explicate the connection, if any, between justification and truth. Reducing the truth connection to the conducing relation is tempting. But that temptation must be resisted.²⁹

References:

- Alston, William P. 1985. "Concepts of Epistemic Justification." *The Monist* 68: 57–89. Reprinted in Alston 1989, 81–114. Page references are to the 1989 reprint.
- . 1988. "The Deontological Conception of Epistemic Justification." *Philosophical Perspectives* 2: 257–99. Reprinted in Alston 1989, 115–52.
- . 1989. *Epistemic Justification: Essays in the Theory of Knowledge*. Ithaca, NY: Cornell University Press.
- . 1991. *Perceiving God: The Epistemology of Religious Experience*. Ithaca, NY: Cornell University Press.
- . 1993a. "Epistemic Desiderata." *Philosophy and Phenomenological Research* 53: 527–51.
- . 1993b. *The Reliability of Sense Perception*. Ithaca, NY: Cornell University Press.
- . 2005. *Beyond "Justification": Dimensions of Epistemic Evaluation*. Ithaca, NY: Cornell University Press.
- Anscombe, G. E. M. 1958. "Modern Moral Philosophy." *Philosophy* 33: 1–19.
- Berker, Selim. Forthcoming. "Epistemic Teleology and the Separateness of Propositions." To appear in *Philosophical Review*. Draft available at <http://www.fas.harvard.edu/~phildept/files/Faculty%20Papers/berker_epist-teleo5.pdf>.
- Broad, C. D. 1930. *Five Types of Ethical Theory*. London: Routledge.
- Conee, Earl, and Richard Feldman. 1998. "The Generality Problem for Reliabilism." *Philosophical Studies* 89: 1–29. Reprinted in Conee and Feldman 2004, 135–65. Page references are to the 2004 reprint.
- and ———. 2004. *Evidentialism: Essays in Epistemology*. Oxford: Oxford University Press.
- and ———. 2008. "Evidence." In *Epistemology: New Essays*, ed. Quentin Smith, 83–104. Oxford: Oxford University Press.
- David, Marian. 2001. "Truth as the Epistemic Goal." In *Knowledge, Truth, and Duty: Essays on Epistemic Justification, Responsibility, and Virtue*, ed. Matthias Steup, 151–69. Oxford: Oxford University Press.
- Feldman, Richard. 2000. "The Ethics of Belief." *Philosophy and Phenomenological Research* 60: 667–95. Reprinted in Conee and Feldman 2004, 166–95. Page references are to the 2004 reprint.
- . 2002. "Epistemological Duties." In *The Oxford Handbook of Epistemology*, ed. Paul K. Moser, 362–84. Oxford: Oxford University Press.

²⁹ This paper was originally written for the 2012 SOFIA Conference on Epistemic Agency in Huatulco, Mexico, where my commentators were Paul Horwich and Joshue Orozco. I also presented versions of the paper at the 2013 MIT-ing of the Minds conference at MIT, where my commentator was Ryan Doody, and at the Workshop on Aims and Norms: Judgment at the University of Southampton, where my commentator was Nick Treanor. Many thanks to all of the participants at these conferences for their extremely helpful questions. In revising the paper for publication I have benefited from conversation and/or correspondence with David Copp, Anil Gupta, Thomas Kelly, Emily McWilliams, Derek Parfit, Shantia Rahimian, Susanna Siegel, and Kurt Sylvan.

- Firth, Roderick. 1978. "The Schneck Lectures." Delivered at the University of Arizona in 1978. First published in Firth 1998, 315–67.
- . 1981. "Epistemic Merit, Intrinsic and Instrumental." *Proceedings and Addresses of the American Philosophical Association* 55: 5–23. Presidential Address delivered at the Annual Eastern Meeting of the American Philosophical Association in December, 1980. Reprinted in Firth 1998, 259–71. Page references are to the 1998 reprint.
- . 1998. *In Defense of Radical Empiricism: Essays and Lectures*, ed. John Troyer. Lanham, MD: Rowman & Littlefield.
- Foley, Richard. 1987. *The Theory of Epistemic Rationality*. Cambridge, MA: Harvard University Press.
- . 1993. *Working Without a Net: A Study of Egocentric Epistemology*. Oxford: Oxford University Press.
- Foot, Philippa. 1985. "Utilitarianism and the Virtues." *Mind* 94: 196–209.
- Goldman, Alvin I. 1979. "What Is Justified Belief?" In *Justification and Knowledge*, ed. George Pappas. Dordrecht: Reidel, 1–23. Reprinted in Goldman 1992. Page references are to the 1992 reprint.
- . 1992. *Liaisons: Philosophy Meets the Cognitive and Social Sciences*. Cambridge, MA: MIT Press.
- . 1999. *Knowledge in a Social World*. Oxford: Oxford University Press.
- . 2001. "The Unity of the Epistemic Virtues." In *Virtue Epistemology: Essays on Epistemic Virtue and Responsibility*, ed. Abrol Fairweather and Linda Zagzebski, 30–48. Oxford: Oxford University Press.
- Grimm, Stephen R. 2009. "Epistemic Normativity." In Haddock, Millar, and Pritchard 2009, 243–64.
- Haddock, Adrian, Alan Millar, and Duncan Pritchard, eds. 2009. *Epistemic Value*. Oxford: Oxford University Press.
- Kagan, Shelley. 2000. "Evaluative Focal Points." In *Morality, Rules, and Consequences: A Critical Reader*, ed. Brad Hooker, Elinor Mason, and Dale E. Miller, 134–55. Lanham, MD: Rowman & Littlefield.
- Kornblith, Hilary. 1993. "Epistemic Normativity." *Synthese* 94: 357–76
- Korsgaard, Christine M. 1993. "The Reasons We Can Share: An Attack on the Distinction Between Agent-Relative and Agent-Neutral Values." *Social Philosophy and Policy* 10: 24–51. Reprinted in Korsgaard 1996, 275–310. Page references are to the 1996 reprint.
- . 1996. *Creating the Kingdom of Ends*. Cambridge: Cambridge University Press.
- Kvanvig, Jonathan L. 2003. *The Value of Knowledge and the Pursuit of Understanding*. Cambridge: Cambridge University Press.
- Muirhead, John H. 1932. *Rule and End in Morals*. London: Oxford University Press.
- Norcross, Alastair. 1997. "Comparing Harms: Headaches and Human Lives." *Philosophy & Public Affairs* 26: 135–67.
- Plantinga, Alvin. 1988. "Positive Epistemic Status and Proper Function." *Philosophical Perspectives* 2: 1–50.
- . 1993. *Warrant and Proper Function*. Oxford: Oxford University Press.
- Pritchard, Duncan. 2007. "Recent Work on Epistemic Value." *American Philosophical Quarterly* 44: 85–110.
- , Alan Millar, and Adrian Haddock. 2011. *The Nature and Value of Knowledge: Three Investigations*. Oxford: Oxford University Press.
- Rabinowicz, Wlodek and Toni Rønnow-Rasmussen 1999. "A Distinction in Value: Intrinsic and for Its Own Sake." *Proceedings of the Aristotelian Society* 100: 33–51.
- Scanlon, T. M. 1998. *What We Owe to Each Other*. Cambridge, MA: Harvard University Press.

- Sosa, Ernest. 1991. *Knowledge in Perspective: Selected Essays in Epistemology*. Cambridge: Cambridge University Press.
- . 2003. “Beyond Internal Foundations to External Virtues.” In Laurence Bonjour and Ernest Sosa, *Epistemic Justification: Internalism vs. Externalism, Foundations vs. Virtue*. Oxford: Blackwell Publishing.
- . 2007. *A Virtue Epistemology: Apt Belief and Reflective Knowledge, Volume 1*. Oxford: Oxford University Press.
- . 2011. *Knowing Full Well*. Princeton, NJ: Princeton University Press.
- Thomson, Judith Jarvis. 1976. “Killing, Letting Die, and the Trolley Problem.” *The Monist* 59: 204–17. Reprinted in Thomson 1986, 78–93. Page references are to the 1986 reprint.
- . 1986. *Rights, Restitution, and Risk: Essays in Moral Theory*, ed. William Parent. Cambridge, MA: Harvard University Press.