# Reliabilism without Epistemic Consequentialism

**Abstract**

This paper argues that reliabilism can plausibly live without epistemic consequentialism, either as part of a non-consequentialist normative theory or as a non-normative account of knowledge on a par with certain accounts of the metaphysics of perception and action. It argues moreover that reliabilism should not be defended as a consequentialist theory. Its most plausible versions are not aptly dubbed 'consequentialist' in any sense that genuinely parallels the dominant sense in ethics. Indeed, there is no strong reason to believe reliabilism was ever seriously intended as a form of epistemic consequentialism. At the heart of its original motivation was a concern about the necessity of *non-accidentality* for knowledge, a concern quite at home in a non-consequentialist or non-normative setting. Reliabilism's connection to epistemic consequentialism was an accretion of the '80s, and a feature of only one of its formulations in that decade.

## 1 Introduction

Recent literature conveys the impression that epistemic consequentialism has long been a dominant view about justified belief.[1] One source of this impression is the prominence of reliabilism in the post-Gettier history of epistemology. While reliabilism has had many critics, it retains many adherents and remains a dominant perspective on the nature of justified belief and knowledge; indeed, it is arguably the leading form of externalism. Accordingly, if reliabilism were a form of epistemic consequentialism, one could reasonably conclude that the latter has been a major force in traditional epistemology since the late '60s and early '70s.

But such thinking would be misguided. Reliabilism in its early formulations was not offered as a consequentialist view, or indeed as a *normative theory* at all. While it took on a relevantly normative formulation in one classic work nearly twenty years after its emergence (viz., Goldman (1986)), this formulation was an accretion, and lacks a claim to dominance. Moreover, reliabilism's apparently consequentialist formulations aren't really analogous to consequentialism in ethics, nor are they best defended in the way consequentialist views are best defended.

Such, at any rate, are some main claims to be defended below. More briefly, my goal is as follows. After defending my historical claims, I will argue that whether or not reliabilism was ever seriously intended as consequentialism, it

[1] See especially Berker (2013a), Ahlstrom-Vij and Dunn (2015), and Dunn (2015).

(C1) can naturally be defended without epistemic consequentialism, and

(C2) should be defended without epistemic consequentialism, since its most plausible forms are not properly regarded as consequentialist: certain restrictions are needed to solve central problems like the clairvoyance and new evil demon problems, and these lack a fundamentally consequentialist rationale.

I will begin with some preliminary clarifications in §2, defend (C1) in §3–4 and (C2) in §5, and take stock and conclude in §6.

## 2 Consequences and the Structure of Normative Theory

Before proceeding to my arguments, I should first explain what kinds of theories I would count as consequentialist, since I will be relying on some assumptions here that are familiar in normative ethics but less familiar in epistemology.

Consequentialism is, I take it, a first-order normative theory, not a meta-normative theory.[2] Since there are different kinds of first-order theory, it is worth considering in more detail what kind consequentialism is. Following Kagan (1992, 1997), we should distinguish *factoral* and *foundational* normative theories. To understand the point of the first kind of theory, note that many normative statuses (rightness, value, justification...) even within a domain (prudential, moral, epistemic...) are a function of various relevant factors. Factoral theorists seek, in Kagan (1997: 17)'s words, to 'articulat[e] these various normative factors, and discover how they interact' to determine statuses like rightness, value and justification. In other words, factoral theories seek to answer the following question of some normative status S in a domain *d*:

(Q1) **The Normative Factors Question**: What factors bear on whether a target of evaluation in *d* has S? What, in other words, are the *S-relevant factors*?

Although factoral theories often posit several factors, they can be monistic. One might, for example, claim that the only factor that bears on whether one's act is prudentially right is its contribution to one's pleasure/pain ratio. This claim can be offered factorally rather than foundationally because one could conceivably go further, adopting a more fundamental but still normative explanation of why one's pleasure/pain ratio is the only prudentially relevant factor. For example, one might defend hedonism about prudential rightness on the basis of

[2] Though see Copp (2009, 2014) for a different approach.

(a) agent-relative act consequentialism about prudential rightness,

(b) a welfarist view about prudential goodness, and

(c) a hedonist account of welfare.

At the foundational level, the theorist offering this position is a consequentialist about prudential rightness and a welfarist about goodness-for. This theorist also believes that welfare is grounded in pleasure, and hence that pleasure is the only factor relevant to prudential rightness.

Despite not being meta-normative positions, foundational normative theories do more than answer (Q1). They seek to answer:

(Q2) **The Normative Foundations Question**: What is the fundamental normative explanation of why the S-relevant factors are S-relevant factors?

Foundational normative ethics hence aims to provide the most fundamental normative explanations of why various normatively relevant factors are the relevant ones. But not all argumentation in normative ethics occurs at this level. Much just concerns which factors are morally relevant. Consider partiality: can the fact that I am in some relationship with X affect how weighty a moral reason I have to benefit X? It is unclear that answering this question requires having already taken a stand on the correct foundational normative theory.[3]

Consequentialism in its most controversial and significant form is, I take it, a foundational normative theory. It must be understood in this way to be inconsistent with paradigmatic non-consequentialist theories. For as I will emphasize again later, the best interpretations of Kantian ethics arguably make it coincide in its predictions with certain forms of consequentialism, but it still differs in *deriving* these predictions from a more fundamental non-consequentialist norm; for example, Wood (1999), Herman (1993: Ch.3; 2007: Ch.11), and Korsgaard (1996: Ch.4) suggest that one can derive the duty to promote the well-being of other people, or to accept rules that would prevent suffering, from more fundamental duties of *respect* or *love* for persons. Indeed, as Dreier (1993) pointed out, arguably any plausible normative theory is extensionally equivalent to some version of consequentialism, but this hardly establishes the

[3]Methodological generalists would say that we must do the foundational work first. But one needn't be a Dancy-style particularist to prefer *methodological* particularism, insisting that we should give a foundational story only after doing the factoral work.

truth of consequentialism. For it remains an open question whether the *explanations* given by the non-consequentialist are preferable. Hence the issue of explanatory priority emerges as the deeper disagreement between consequentialists and non-consequentialists.[4] This fact will prove important in my argument in §5.

In its central versions in ethics, consequentialism seeks a certain sort of normative explanation of facts about rightness and related normative properties in terms of value: in particular, an *instrumental* explanation that reveals a *means-end* link between rightness and the promotion of value.[5] As Berker (2013a: 343) notes, this qualification on the form of the consequentialist's value-based explanation is important. There are theories that are value-based but not consequentialist, like virtue ethics and Herman (1993) and Wood (1999)'s interpretations of Kantian ethics, which suggest that the value of humanity or the value of the good will is normatively fundamental, but which insist that the fundamentally proper response to this value is to respect it. In its *direct* form, consequentialism about right action upholds the following claim:

> **Direct Consequentialism about Right Action**: The fundamental normative explanation of why right actions are right is that they are conducive to final value.

In its broadest *indirect* form, consequentialism about right action upholds the following claim:

> **Indirect Consequentialism about Right Action**: The fundamental normative explanation of why right actions are right is that they are produced by something that is conducive to final value,
>
> – where this 'something' might be the *general acceptance of a certain rule* (rule consequentialism) or the *operation of a certain motive* (motive consequentialism).

Now, it is important to add two qualifications to these preliminary glosses. Firstly, it is worth stressing that consequentialists of both sorts have for crucial reasons understood value-conduciveness in terms of both the proximal and distal consequences of an act, rule, or motive. Of course, to solve the problem of demandingness, a switch from *actual* to *expected* consequentialism might be needed in theorizing about *decision procedures* rather than *right-making characteristics*,[6]

[4] See esp. Scanlon (1998: Ch.2) for this point. As Scanlon argues, consequentialism rests on a misunderstanding of what intrinsic value fundamentally demands, by treating it as fundamentally 'to be produced' despite the fact that many plausible examples of intrinsic values call fundamentally for respect or appreciation (e.g., personhood, friendship, art), and only derivatively for production.

[5] Here 'means' is used broadly to encompass not just efficient causal means, but also constitutive means (e.g., pleasure might be a means to well-being in virtue of *constituting* well-being, on hedonist views).

[6] See Bales (1971) and Railton (1984) for this distinction.

where the latter are the real focus of foundational theorizing. But making this switch doesn't involve claiming that only the expected proximal consequences matter. One can easily have expectations about some distal consequences. And such a theory would anyway be strange: if what fundamentally matters is bringing about the good, why would one ignore the potentially massive outweighing of proximal goodness by distal badness? Not even a desire to appease the time-biases of ordinary agents would justify such a restriction, since consequences can be distal while still being short-term and expected.

No direct consequentialist of either a maximizing or a satisficing sort would claim that moving one's finger is right when it would predictably lead to pointless near-term though not immediate disaster. For example, the fact that the only proximal consequence was the pressing of the pretty red button (which the agent enjoyed) is no justification if the distal consequences involve nuclear war. Similarly, no indirect consequentialist of either a maximizing or satisficing sort would claim that it could be right to act in accordance with a rule whose general acceptance would predictably lead to pointless near-term though not immediate disaster. Nor will *agent-relative* versions of consequentialism embrace these implications except for agents with eccentric utility functions. Such agents will still take distal consequences into account: they will just give them less weight. Other kinds of discrimination may appear in factoral theorizing or theorizing about decision procedures, but discrimination always invites explanation.[7] From a foundational consequentialist perspective, bad consequences of the relevant object of evaluation cannot be ignored until they are outweighed by good ones.

Littlejohn (2018: 25) offers a helpful way of capturing this point without excluding certain paradigmatic consequentialist theories. He suggests that consequentialists are committed to the *totalizing* assumption that the total consequences of whatever is directly evaluated by the theory must figure in the determination of rightness. It is left open to consequentialists to decide how to assign (dis)value to these many consequences. But the only way to practically exclude distal consequences would be by assigning neutral value to them. As Littlejohn emphasizes—following Foot (1985)—this totalizing assumption is essential for capturing the intuitive appeal of consequentialism: 'the reason that consequentialism seems so compelling is

[7] Indeed, Sidgwick (1907: 379–387)—the most consistent of the classical utilitarians—took this point to reflect a self-evident principle of rationality.

that when it comes to the good, the better, and the best, it is irrational to prefer some acknowledged lesser good to one that is greater' (25). If one agrees that X's distal badness would be much worse than its proximal goodness, it would seem one prefers the acknowledged lesser good in preferring X (unless there are no alternatives with better distal consequences). Again, one could imagine an eccentric agent-relative view that was indifferent to distally produced suffering while assigning disvalue to proximally produced suffering. An agent who did what this view required wouldn't be irrational in the narrowest sense (i.e., akratic). But the view would achieve this prediction not by claiming that distal consequences are no part of the calculation, but rather by taking them into account and assigning them neutral value.

This last point is important. Alhstrom-Vij and Dunn (2018: 1) suggest that ethical egoism is not a totalizing theory, because it only pays attention to the good that is brought about for the agent.[8] But ethical egoism does pay attention to the total consequences. It just evaluates those consequences solely in terms of their goodness for the agent. It is not that some consequences are ignored: instead, *only certain values* are brought to bear in the overall evaluation of the total consequences. So, contrary to what they suggest, there is no familiar version of consequentialism in ethics that rejects the totalizing requirement. Such a theory would abandon what is arguably the most attractive feature of consequentialism, and the one needed to explain its appeal. So I think it would be better, as Ahlstrom-Vij and Dunn temporarily acknowledge in summarizing Wedgwood (2018)'s view, to call a view that rejects the totalizing requirement *teleological* but not consequentialist.[9]

My second qualification stems from a further point that Littlejohn (2018) emphasizes.[10] As he observes, following Pettit (1991: 230-33; 1997: 125-30) and Scanlon (1998: 79-80), there is a deeper reason why consequentialists take norms to be explained by means of *conducivity* to

[8]They seem to suggest this when they contrast act utilitarianism with ethical egoism: 'Notice several other commitments of act utilitarianism. First, it maintains that rightness is determined not just by *some* of the consequences of actions but by the *total* set of consequences.... Not all views that have been said to be instances of consequentialism share these features. We can, for example, imagine a form of ethical egoism, where actions are right when maximizing the good *for the agent herself*' (1).

[9]See Alhstrom-Vij and Dunn (2018: 3). Wedgwood (2018: 90) is clear on the importance of the totalizing requirement for consequentialists: '[C]onsequentialism is best understood as the doctrine that the value or normative status of the relevant states of affairs (whether these states of affairs be *acts*, or *beliefs*, or anything else) is derivative from the value of these states of affairs' *total consequences*.'

[10]This point is also briefly noted by Berker (2013a: 343) but not sufficiently considered in his case for the pervasiveness of epistemic consequentialism.

fundamental value. The reason is that they understand fundamental value in a certain way: what it is to be fundamentally valuable is to be something that is 'to be promoted' for its own sake (where 'promote' means *bring about*, not *advertise* or *campaign for*). Value-based theories can reject this assumption, but they thereby cease to be consequentialist.

With these two qualifications in mind, I think it is most faithful to the ethics literature to restate direct and indirect ethical consequentialism in the following ways:

> **Direct Consequentialism about Right Action**: The fundamental normative explanation of why right actions are right is that their total consequences sufficiently promote fundamental value, where 'fundamental value' just is value that is to be promoted for its own sake.[11]
>
> **Indirect Consequentialism about Right Action**: The fundamental normative explanation of why right actions are right is that they are produced by something whose total consequences sufficiently promote fundamental value,
>
> – where this 'something' might be the *general acceptance of a rule* (rule consequentialism) or the *operation of a motive* (motive consequentialism).

We can now see what it would take for a theory of justified belief to be consequentialist in a way that yields a faithful analogy with ethics. By analogy with the foregoing formulations of consequentialism about right action, it could take a direct or indirect form:

> **Direct Consequentialism about Justified Belief**: The fundamental normative explanation of why justified beliefs are justified is that their total consequences sufficiently promote fundamental epistemic value, where 'fundamental epistemic value' just is value that ought epistemically to be promoted as an end.
>
> **Indirect Consequentialism about Justified Belief**: The fundamental normative explanation of why justified beliefs are justified is that they are produced by something whose total consequences sufficiently promote fundamental epistemic value,
>
> – where this 'something' might be the *acceptance of a certain epistemic rule* or the *operation of a cognitive process or disposition*.

To the extent that consequentialism has plausibly had any significant history in epistemology, it is in the second of these forms. But as we will see, there are some striking disanalogies between indirect consequentialist theories in ethics and the epistemological theories that some

[11] Here and below, 'sufficiently' isn't meant to exclude maximizing consequentialism, which could be understood as imposing a perfectionist bar on being good enough.

have claimed to be indirectly consequentialist. In what follows, I will assume that if the analogy is strained too far—so that views resembling paradigmatic non-consequentialist theories end up being called 'consequentialist'—it will not be worthwhile to regard them as genuinely consequentialist. For example, I will assume that a theory is at best dubiously worthy of the name 'consequentialist' if it embraces at the fundamental level *restrictions* or *side-constraints* on promoting the good, or if it insists that value is fundamentally to be *respected*, not promoted. Because of its role in sustaining the appeal of consequentialism, I will also assume that the totalizing assumption is essential to consequentialism, where adjustments can be made to the underlying value theory to mimic non-totalizing theories.

These assumptions do not involve unfair deck-stacking. If epistemologists claiming the label 'epistemic consequentialism' reject the analogues of claims that are essential for motivating consequentialism in contrast to the various non-consequentialist theories, their use of the label threatens to become misleading at best, and merely homonymous at worst. The characterization I've given above captures what other central players in normative ethics agree to be key features of the view when understood as a foundational normative theory (which is the only format in which it must disagree with Kantianism or virtue ethics). This characterization is also embraced by Berker (2013a), though I believe he is too fast in judging theories to conform to it. As I will discuss in considering Goldman's reliabilism, my characterization does conflict in one respect with the ultimate characterization given by Alhstrom-Vij and Dunn (2014, forthcoming), though they begin with a characterization like my first. But their characterization jeopardizes the analogy with ethics. Moreover, the reasoning that first led them to this characterization in their (2014) involved assuming that reliabilism is a form of epistemic consequentialism. While they treat this assumption as non-trivial in Alhstrom-Vij and Dunn (forthcoming), their willingness to abandon paradigmatic features of the consequentialism/non-consequentialism contrast in ethics is not defensible.

Now, when consequentialism is understood as a foundational normative theory, it is possible for some views to *appear* consequentialist from a certain angle without *really being* consequentialist. Note that factoral theorizing will plausibly *underdetermine* foundational theorizing even if our factoral inventory is highly constrained. This underdetermination probably won't leave

the foundational options wide open. But to expect no underdetermination between at least one consequentialist theory and at least one non-consequentialist theory seems optimistic. If life turns out even a little hard, we will find ourselves choosing between a consequentialist theory and a non-consequentialist theory that from a distance may sound consequentialist because it agrees at the factoral level with some consequentialisms. This fact plays a central role below. The evidence for some forms of reliabilism may look like evidence for epistemic consequentialism, since it is also predicted by epistemic consequentialism. But the best foundational explanation of this evidence and the best underpinning for reliabilism is non-consequentialist.

There is a final assumption I should make explicit, which should be obvious but sometimes seems neglected in epistemology. The assumption is that consequentialism is a *normative theory*, in a broad sense of 'normative' that includes both theories of rational decision procedures and theories of right-making characteristics. Note that there are all manner of theories in philosophy that ground certain kinds of facts about an X in facts about X's effects, or in facts about the effects of something else. Since effects are consequences, these theories ground certain kinds of facts in facts about consequences. Many of these theories are obviously not in the same ballpark as consequentialism. Consider, for example:

- *Causal theories of property identity*—e.g., the view that F and G are identical iff F and G confer the same causal powers on all particulars that instantiate them.
- *Causal theories of perception*—e.g., the (admittedly oversimplified) view that a visual experience is a seeing of x iff x causes that visual experience.
- *Causal theories of intentional action*—e.g., the (admittedly oversimplified) view that a bodily movement is an intentional action iff it is caused by an intention.
- *Response-dependent theories of color*—e.g., the view that X is red iff X causes red-experiences in observers in normal conditions.

It might be heuristically useful to analogize these theories with forms of consequentialism. But it would be a mistake to frame this analogy as the discovery that consequentialism is a popular view in the metaphysics of color, property identity, perception, and action. Such framing would be confused partly because color, perception, and action seem non-normative. As a result, these causal theories aren't in the same business as consequentialist theories. But even if there were some sense in which color and action, say, were normative, it still wouldn't follow that the

explanations given by causal theories of color and action would be normative explanations. For the relation between perception and its causal constitution as well as between action and its causal constitution is not normative, while the relation between an act's rightness and what makes it right in first-order terms is normative.

These points matter in assessing whether reliabilism is rightly viewed as a form of consequentialism. Reliabilism in its earliest formulations was a theory about the nature of knowledge. Knowledge might just be a generic factive mental state. Even if this state were in some sense normative, a causal theory of its constitution analogous to the causal theory of perception would not be a consequentialist theory. For the constitution relation in play needn't be a normative relation of the kind in play in first-order ethics.

# 3 Reliabilism's Life without Consequentialism (and Occasional Dalliance with It)

I turn now to a historical argument for (C1). Here in outline are the main points this section will make. Reliabilism lived for more than ten years as a non-normative theory. It was treated in Goldman (1979) as similar in aim to various first-order ethical theories. But Goldman (1979)'s theory could be endorsed factorally or foundationally, and if factorally endorsed it could be embedded within a non-consequentialist foundational view—e.g., virtue epistemology, which was supposed to be analogous to one main alternative to consequentialism in ethics.[12]

While Goldman (1986) later proposed an analogy between his project and the rule consequentialist's, this was just one moment in the history of reliabilism, and much of Goldman's later work didn't build on the analogy. Anyway, as I will suggest, Goldman (1979, 1986)'s views aren't analogous to indirect consequentialism: the analogy is flawed in ways that cannot be ignored if we are serious about our comparisons with ethics. Finally, other reliabilists have continued to hold a more traditional, non-normative form of reliabilism. So although Goldman's contributions are preeminent, one shouldn't forget that reliabilism receives notable treatment in other hands.

[12]Goldman (1993: 274) suggests this framing.

I turn to document these points in more detail.

## 3.1 Pre-1979 Reliabilism: Non-Accidentality and the Constitution of Knowledge

Reliabilism started out as a theory of knowledge, not a theory of justification.[13] What the early versions of reliabilism in Armstrong (1968, 1971), Unger (1968), Dretske (1971), and Goldman (1975) have in common is an attempt to capture the non-accidentality condition on knowledge. While it is standard after Pritchard (2012) to distinguish anti-luck conditions from ability conditions, and reliability conditions are arguably best understood as versions of the latter, this distinction is new.[14] Instead of taking it to show that non-accidentality approaches are not necessarily reliabilist, we could take it as an invitation to distinguish different types of reliabilist theory—say, process reliabilism on the one hand and 'counterfactual reliabilism', to use Vogel (2000)'s term, on the other. Indeed, there is plausibly a sense of 'accident' for which it is necessarily true that it's no accident when reliable thinkers' beliefs turn out true. This fact remains even if environmental luck cases show there to be some other condition on knowledge, negatively connected to another sense of 'accident'.[15]

If one is a reliabilist of the classic stripe, it would be coherent for one to either withhold belief in epistemic consequentialism or reject it. The epistemic consequentialist would derive her reliability condition on knowledge from three things: the assumption that knowledge implies justified belief, an indirect epistemic consequentialist account of justified belief, and a veritist epistemic axiology. The traditional reliabilist needn't accept the reliability condition for such reasons. She can be a reliabilist simply because she thinks knowledge is non-accidentally accurate belief, and reliability provides the best unpacking of such non-accidentality.

There is no necessary tie between embracing a view that sees non-accidentality and the

[13]Ramsey (1931b: 193–198) provides an interesting exception in giving a reliabilist account of reasonable belief, and there is evidence that he was an epistemic consequentialist (see esp. p.196). But as Goldman and Beddor (2015) observe, Ramsey's early reliabilism 'attracted no attention at the time and apparently did not influence reliability theories of the 1960s, 70s, or 80s.'

[14]Goldman and Beddor (2015) describe Goldman (1976)'s response to the fake barns case as 'counterfactual reliabilism', however.

[15]But see Millar (2010) and Littlejohn (2014) for a derivation of the anti-luck condition from the ability condition.

main or even the only thing one must add to true belief to get knowledge and embracing epistemic consequentialism. Indeed, a close parallel in ethics to early reliabilism's emphasis on non-accidentality is provided by interpretations of Kant that see *non-accidental conformity to moral law* as the defining feature of moral worth, and what is missing in examples like the shopkeeper case, which Kant (1785/2012: 4.398) describes as a case of 'fortunate' (i.e., accidental) rightness.[16] Indeed, the best recent work explaining the thought that moral worth is a matter of doing the right thing for the right reasons has itself been inspired by the 'AAA' pattern from virtue epistemology: according to Mantel (2018) and Lord (2018), for example, an act only has full moral worth if it is right *as a manifestation of reasons-sensitive dispositions*, which both suggest is a kind of 'aptness' in Sosa's sense. It is not enough, according to these theorists, to say that moral worth consists in the 'coinciding' of right-making reasons with one's motivating reasons (as Markovits (2010) suggested): the coinciding itself must manifest reasons-sensitive dispositions.

One can also be a reliabilist of the classic stripe without offering reliabilism as a normative theory of any kind. Armstrong (1968: Ch.9)'s first presentation of his reliabilism about knowledge appeared in a book on the metaphysics of mind, where he treated knowledge a mental state and saw perceiving to be a special case of knowing (see (1968: Ch.10) and cf. (1961)). On this kind of view, one could hold that knowledge is no more normative than the factive mental states that are its determinates.[17] If so, the condition of non-accidental arrival at the truth will just be a generalization of the condition of non-accidental veridicality on seeing that p. It no less a mistake to claim that this kind of view is an example of consequentialism in epistemology than it is to say that the causal theory of properties is a version of consequentialism in metaphysics.

It is worth adding that even if knowledge were clearly a normative relation in the same sense in which paradigmatically properties like justification and obligation are normative, it still wouldn't follow that a metaphysical account of the nature of knowledge is a normative theory in the sense in which consequentialism is a normative theory. Ethical theorists commonly hold that giving an account of the *constitution* of something normative is different from giving a

[16]For some different versions of this interpretation, see Herman (1981), Benson (1987), Baron (1995), and Stratton-Lake (2000).

[17]For further defense of this view, see Sylvan (2018).

first-order normative theory. First-order normative theories do not give *analyses* of normative properties—a task reserved for meta-ethics—but rather substantive theories that explain in a distinctive normative way why certain acts or attitudes have these properties.

## 3.2 1979 Reliabilism

It must be acknowledged that reliabilism did in the hands of Goldman (1979) become a normative theory in a broad sense. Goldman was explicit from the first page:

> The term 'justified', I presume, is an evaluative term, a term of appraisal. Any correct definition or synonym of it would also feature evaluative terms. I assume that such definitions or synonyms might be given, but I am not interested in them. I want a set of substantive conditions that specify when a belief is justified. This might be defined in other ethical terms or phrases, a task appropriate to meta-ethics. The task of normative ethics, by contrast, is to state substantive conditions for the rightness of actions. Normative ethics tries to specify non-ethical conditions that determine when an action is right. [...] Analogously, I want a theory of justified belief to specify in non-epistemic terms when a belief is justified.[18]

On the same page, Goldman mentions act utilitarianism as an example of a theory that gives 'substantive conditions' for rightness. But he doesn't suggest that his view is analogous to utilitarianism in more than the general sense clarified in the above passage. We are not getting the overt modelling on consequentialism here that we get in Goldman (1986).

Still, one might think that the normative theory Goldman (1979)'s view best resembles is a monistic indirect consequentialism—say, rule utilitarianism or, closer yet, motive utilitarianism. For this reason, one might not find it misleading to say that indirect epistemic consequentialism received its first detailed formulation in Goldman (1979). But even this claim lacks plausibility. Goldman (1979)'s theory is in key ways not parallel to indirect consequentialism.

There are at least three features of Goldman (1979)'s account that support this verdict, on which I'll comment briefly in outline and then at greater length:

A. The account counts only certain highly restricted outputs of processes in the measure of their reliability, which is not what one would expect in a theory analogous to any familiar consequentialism in ethics.[19]

[18] Goldman (1979: 1).
[19] Goldman (1979: 12–13).

B. Goldman's account of defeat has no analogue in any familiar consequentialism in ethics. While it is a more purely reliabilist account of defeat than the one found in some of his later work (see, e.g., Goldman (1986) and Goldman (2011)), it is not thereby more consequentialist, but less so.

C. The asymmetry the account yields in the conditions for justified basic belief vs. justified non-basic belief has no real analogue in consequentialism in ethics.[20] The closest analogue of the restrictions Goldman places on non-basic beliefs are *side-constraints*, which are paradigmatically non-consequentialist.

Let me now elaborate on each, taking (A) first.

While the reliability of a process is determined by looking at certain properties of its consequences, only a highly restricted class of consequences matters—viz., only the 'immediate' outputs of the process, which Goldman (1979: 13–14) takes into account in his clauses (6A) and (6B). This restriction lacks precedent in the consequentialist tradition in ethics. It conflicts with the totalizing requirement mentioned earlier as essential for explaining the intuitive appeal of consequentialism. Direct consequentialists take all consequences of an act to be relevant, and indirect consequentialists take all consequences of whatever they directly evaluate in consequentialist terms—rules or motives—into account. Moreover, while one could imagine a factoral theory that restricted consequences to immediate ones, it is hard to tolerate such a restriction as a piece of fundamental normative theorizing that is genuinely consequentialist in spirit. Why should the proximal/distal distinction have such significance? The non-consequentialist, of course, has an answer that would derive this restriction from some more fundamentally appealing restriction. The consequentialist does not, unless distal consequences are considered but given less weight in the associated value theory.

Now, Ahlstrom-Vij and Dunn (2014) try to use the presence of the immediacy restriction in reliabilism to answer Berker (2013a)'s argument against epistemic consequentialism, insisting that reliabilism is both a direct and an indirect consequentialist theory. But they begged the question there by assuming without argument that reliabilism is a form of epistemic consequentialism,[21] and by using 'direct' in a sense not used in the consequentialist literature in ethics. Because they assumed there that reliabilism is a form of consequentialism, they assumed that the immediacy restriction is available to the consequentialist in virtue of its availability to

[20] Goldman (1979: 13–14).

[21] To be fair, Berker (2013) also paints reliabilists as epistemic consequentialists.

the reliabilist. But it is *not* available to reliabilists for consequentialist reasons, if one uses 'consequentialist' in a sense that parallels ethics. While Ahlstrom-Vij and Dunn (forthcoming) now acknowledge that it is an open question whether reliabilism is a kind of consequentialism, they end up admitting that Goldman's view contains the epistemic analogue of side-constraints. They also do not give a precedent in the consequentialist literature for the immediacy restriction: their example of ethical egoism is problematic for reasons explained earlier.

A third feature of Goldman's account that is hard to derive from consequentialist foundations is its theory of defeat. Of course, the very idea of defeat is not foreign to consequentialism: an action can be right in one circumstance but wrong in another if it has different consequences in the two. Examined from a distance, adding or subtracting consequences might resemble adding or subtracting information. But the particular way in which reliabilists model defeat is not what one would expect if reliabilism were given consequentialist foundations. Goldman (1979)'s account of defeat is especially interesting to consider, since it is his only purely reliabilist account of defeat. He moved in (1986) to allow beliefs and experiences to be defeaters irrespective of ancestry, and continued in (2011) to give experiences this role.

Goldman (1979)'s account of defeat is that a belief produced by a reliable process has *prima facie* but not *ultima facie* justification if there is an alternative reliable process available to the believer that would have resulted in a different doxastic attitude (disbelief or suspension). There is no direct or indirect consequentialist analogue of this account of defeat. There is no direct consequentialist analogue of this conception of defeat, since direct consequentialism is forward-looking and the general value of the propensities of act-producing processes are swamped in rightness evaluation by the particular consequences of the token act. Perhaps there is a vague analogy between the relevance of alternative processes and the relevance of alternative strategies open to the agent at the time of action. But note that if two strategies that recommend different acts would have equally good consequences, act consequentialism permits either; by contrast, if one reliable process would recommend belief and an equally reliable process would recommend suspension of judgment, process reliabilism doesn't permit either belief (and rightly so).[22]

[22]Cf. Littlejohn (2012: 77), who notes that epistemic consequentialism fails to secure this prediction but maligns Goldman because he assumes Goldman is a consequentialist.

For a related reason, there is also no indirect consequentialist analogue of Goldman (1979)'s conception of defeat. Consider rule consequentialism first. If two rules that recommend different acts A and A* in some circumstance C are such that their acceptance would have equally good consequences, or if two rules are equally reliable guides to promoting the good, it is permissible to do either A or A* in C. Something similar goes for motive consequentialism: if two motives are equally reliable guides to promoting the good, or are such that acting on either would tend to produce similar amounts of goodness, then acting on either is permissible. But again, if one reliable process would recommend belief and an equally reliable process would recommend suspension of judgment, process reliabilism (rightly) doesn't permit both.

Curiously, the less purely reliabilist stories Goldman later tells about defeat are more similar to what consequentialist foundations would encourage. Note that the consequentialist can help herself to the structural distinction between subjective and objective rightness. She might hold that justified action requires subjective rightness, and hence must have good expected consequences, or be in accordance with a rule whose acceptance would have good expected consequences. Goldman's later treatment of beliefs and experiences as potential defeaters might be read in a similar way: if some beliefs or experiences can influence expected reliability, they can thereby impact epistemic justification. But these proposals about defeat in epistemology are often regarded as *ad hoc*: as a thoroughgoing externalist view, reliabilism only explains why *actual* reliability matters, not perceived or apparent reliability.[23]

A final feature of Goldman (1979)'s account that doesn't straightforwardly follow from consequentialist foundations is its (modest) foundationalist structure, and specifically the asymmetry it yields between the justification conditions for basic and non-basic beliefs. Of particular interest is the requirement in Goldman's recursive account that non-basic beliefs not be based on unjustified beliefs on pain of inheriting this unjustifiedness. The analogue of this requirement in the practical case would be a requirement that non-basic actions not be based on wrongful actions. Now consider the practical basing relation of immediate interest to the consequentialist, which is the means-end relation. If consequentialism were strongly parallel to Goldman's reliabilism, we would expect the consequentialist to claim that a non-basic action

[23] See, e.g., Greco (2010: 158)

cannot be right if it is performed via wrongful means.

But consequentialists shouldn't like this claim. While a consequentialist can allow that a killing of an innocent person is *pro tanto* bad and *prima facie* wrong in virtue of realizing an intrinsically bad state of affairs, a killing can become right *ultima facie* if its effects are good enough. While some consequentialists set high thresholds, they will eventually allow tradeoffs. Otherwise they admit side-constraints, and as I will explain later, embracing side-constraints in a foundational theory makes one a non-consequentialist. The thinking here is the reverse of what seems sensible in the epistemic case. The all-things-considered deontic evaluation of the basis belief is prior to the evaluation of the based belief, while the all-things-considered deontic evaluation of the means action is posterior to the evaluation of the goal action.

Note again that I am not saying that there won't be heuristically interesting analogies between the foundationalist's structure for justified beliefs and some normative structures that consequentialists don't frown upon. There is, for example, an interesting analogy between regress arguments for foundationalism and regress arguments for intrinsic value. But what we need is not some indirect analogy between a theory of epistemic justification and a theory of value, but rather an analogy between a theory of epistemic justification and a theory of *right action*. For there to be an analogy strong enough to say that the same view appears in both domains, we need more than rough resemblances between unlike normative categories. What we need are parallels between like normative categories—i.e., justified belief and justified action.

Goldman did draw an analogy between reliabilism and rule consquentialism in *Epistemology and Cognition*. But despite this fact, many of the disanalogies to which I've drawn attention remain: Goldman (1986) keeps similar restrictions on the kinds of processes and outputs that matter, and preserves an asymmetry between basic and non-basic beliefs like that in Goldman (1979). Given those remaining disanalogies, Goldman's view is more similar in the end to moderate deontology.[24]

[24]While that claim might sound strange, don't forget that deontologists agree that there is a strong reason to promote the good: they just think this reason can be defeated if the only way to promote the good would be by violating restrictions.

## 3.3 Non-Goldmanian '80s Reliabilism

*Epistemology and Cognition* was not the only memorable event for reliabilism in the '80s. A different reliabilism—indicator reliabilism—also flourished in that decade. While the indicator reliabilism of the time was overshadowed by Goldman's approach, insights from '80s indicator reliabilist theories appear in Goldman's recent views (e.g., Goldman (2011)). As I'll argue shortly, if indicator reliabilist theories are counted as consequentialist, then views in ethics that aren't consequentialist will be wrongly classified as consequentialist.

Besides indicator reliabilism, the '80s also featured reliabilist theories of knowledge that—like Goldman (1967)'s causal theory—denied that justification is necessary for knowledge and are consistent with thinking of knowledge as no more normative than factive mental states like seeing that p.[25] So while *Epistemology and Cognition* was a classic moment in the history of reliabilism, it would be wrong to think that reliabilism took a definite turn toward consequentialism in the '80s.

Let's further consider indicator reliabilism, and then briefly consider the clearer case of '80s reliabilism about knowledge. Foreshadowed by some externalist accounts of knowledge in the '70s,[26] indicator reliabilism about justification received subtly different formulations in Swain (1981a, 1981b, 1985) and Alston (1985, 1988). Alston's account is more intuitive, so I'll take it as the paradigm. Alston's account is a two-level account that first identifies justified beliefs with beliefs based on adequate reasons, and then identifies adequate reasons with reliable indicators. I do not think that the structure of this theory automatically recommends a consequentialist interpretation. If incorporating reliability in an account of justification in this way were sufficient for making the account consequentialist, then ethical theories that are not consequentialist will be wrongly classified as consequentialist.

I will take an appealing combination of views in ethics to illustrate this point. I find it attractive to combine (a) an account of practical justification in terms of possessed objective reasons

[25] I have especially in mind Dretske (1981). While Foley (1987) famously defended an internalist theory of rational belief, he also recommended divorcing the theory of knowledge and the theory of justified belief and indicated his sympathies for a reliabilist theory of knowledge.

[26] See Dretske (1971) and Armstrong (1973)'s 'thermometer view', a name unwittingly hearkening back to an indicator reliabilist theory of meaning by the same name that Price (1953: 185-97) attacked. Sellars (1956: §VII.31) approvingly alludes to Price's attack just before anticipating and attacking the indicator reliabilist account of knowledge in §VIII.

with (b) an account of objective reasons for action that analyses them in terms of objective evidence and rightness *a la* Thomson (2008) and Star (2015). This account is structurally like Alston's: (a) is parallel to the first layer of his view, and (b) is parallel to the second layer, with objective evidence being naturally construed as the same thing as a reliable indicator.

A deontologist can hold (a) and (b). Thomson is a paradigmatic deontologist, and she accepts (b) and could accept (a). Star contrasts his view with indirect consequentialism but embraces (a) and (b). Consistency is possible here for many reasons, one of which is that rightness can be understood as conformity to deontological standard(s). The same goes for epistemology. The truth norm could be understood as a deontological standard of the same sort. Just as (a) and (b) are compatible with moral deontology, so Alston's view is compatible with a truth first epistemic deontology, which I'll describe further in the next section.

The more fundamental point here is the simple one emphasized earlier that one intuitively attractive theory can be underpinned by different normative foundations. Indicator reliabilism can be underpinned by different conceptions of how truth governs belief. A non-consequentialist way to think of how truth governs belief is to think that there is an objective norm of correctness for belief—it is correct to believe p only if p—and that good epistemic reasons for belief are objective evidence that the belief that p would be in conformity with this norm.

Of course, one might frame indicator reliabilism in a gerrymandered consequentialist way. One could take the fundamental epistemic norm not to be a deontological norm of correctness but rather a consequentialist principle that says 'Promote a high ratio of true to false beliefs!', and then preserve the epistemic analogues of (a) and (b). But this fact merely shows that a plausible non-consequentialist view in epistemology can be extensionally equivalent to a gerrymandered consequentialist view. The case-based evidence for indicator reliabilism doesn't decide between non-consequentialist and gerrymandered consequentialist framings.

One might try to bracket indicator reliabilism on the grounds that it is not a purely reliabilist view, but rather a hybrid with internalist and externalist elements. It was intended as such a hybrid in Alston (1988). But the mere fact that a view has internalist and externalist elements is not a sufficient reason for disqualifying it as a form of reliabilism. Goldman (1986)'s theory is reliabilist if any theory is. But it allows internalist factors to defeat externalist *prima*

*facie* justification. There is admittedly an important distinction between views on which reliability is merely necessary for justification and views on which it is necessary and sufficient for justification. But the former views are full-fledged externalisms if the reliability required is reliability vis-à-vis external world propositions. In any case, the most popular versions of reliabilism about justification don't hold that reliable formation is necessary and sufficient for justified belief across the board. Even Lyons (2009)—perhaps the most 'mad-dog' reliabilist around—rejects this view for non-basic beliefs.

So much for how indicator reliabilism could be given non-consequentialist foundations. I'll conclude with a brief mention of '80s reliabilism about knowledge. Dretske's *Knowledge and the Flow of Information* was as important an event for externalist epistemology as Goldman's *Epistemology and Cognition*. Dretske (1981: Ch.4) continued in the footsteps of the early Goldman, denying that knowledge should be analyzed in terms of justification, and instead offering a synthesis of earlier causal accounts of knowledge with insights from the indicator reliabilist tradition. This view is a further reminder that reliabilism can be understood as a non-normative theory, and hence not a consequentialist theory.

## 3.4 Virtue Reliabilism

Another theory often classified as a version of reliabilism is the virtue epistemology of Sosa (1991, 2007, 2015) and Greco (2010); indeed, when this view was first introduced at the end of 'The Raft and the Pyramid', the section title used to introduce it was 'Reliabilism: An Ethics of Moral Virtues and an Epistemology of Intellectual Virtues'. But this view is not a version of epistemic consequentialism.

If this view is intended not only as an alternative to other forms of reliabilism, but as genuinely analogous to virtue ethics (hence earning its name), it is best read as having two levels, one foundational and the other factoral. At the foundational level, the account is the epistemic analogue of virtue ethics. As a result, it is analogous to an alternative to the other major foundational options, which are epistemic consequentialism and deontology. At the factoral level, the account is extensionally equivalent to a form of reliabilism in which person-level dispositions worthy of the title 'competences' rather than potentially subpersonal processes are the

key justifiers. It is then a first-order matter that the normative properties identified as crucial at the foundational level—viz., virtues—line up with the non-normatively specifiable reliable dispositions which do the work at the factoral level.

If this is the correct interpretation of reliabilist virtue epistemology, it is not a form of epistemic consequentialism. Now, some critics—e.g., Dancy (1995) and Zagzebski (1996)—have insisted that the conception of virtue with which Sosa and Greco work is so thin that the account turns out to be consequentialism in disguise. But this criticism is not persuasive, especially if one examines the most recent articulations of the view. It is natural to see Sosa and Greco as inspired by the Aristotelian thought that the fundamental normative property is attributive goodness, which has performance normativity as a special case. It is natural to see them as holding in addition that as a substantive matter, manifestations of unreflective dispositions make for good performances in this sense. Perhaps this view will be extensionally equivalent to something like motive consequentialism, but the reason for the 'thinness' of its conception of virtue is not some underlying consequentialism, but rather that this is just plausible on its face when one reflects on some paradigm cases of knowledge.

## 3.5 Given the History, Why Do Some Deem Reliabilism Consequentialist?

I hope I've shown that there is little evidence for the claim that reliabilism is consequentialist in a sense that genuinely parallels consequentialism in ethics. Given the historical record, one might be left wondering why recent discussions of epistemic consequentialism have been so quick to assume that reliabilists are consequentialists. Before turning to some non-historical arguments for my claims, I will address this lingering question in a way that does justice to the recent literature while still maintaining that it errs.

A side-by-side comparison of process reliabilism and a simple formulation of epistemic consequentialism makes it easy to see why Berker (2013a, 2013b, 2015) and Ahlstrom-Vij and Dunn (2014, forthcoming) thought it clear that reliabilism is a kind of epistemic consequentialism. Consider Ahlstrom-Vij and Dunn (2014)'s gloss on epistemic consequentialism and Berker (2013b)'s gloss on consequentialism *simpliciter*:

> Epistemic consequentialism is the idea that the epistemically right (e.g., the justified) is to be understood in terms of conduciveness to the epistemic good (e.g., true belief). (Ahlstrom-Vij and Dunn 2014: 541)

> 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). (Berker 2013b: 264)

Now compare a simplified statement of process reliabilism from Goldman (2012):

> [T]he main proposal of 'What Is Justified Belief?' was that a belief's justifiedness is fixed by the reliability of the process or processes that cause it, where (as a first approximation) degree of reliability consists in the proportion of beliefs produced by the process that are true. Justification-conferring processes are ones with a high truth-ratio. (Goldman 2012: 73)

It seems obvious enough why process reliabilism might appear to be a version of consequentialism according to Ahlstrom-Vij, Dunn, and Berker. Process reliabilism in the simple formulation does say that whether a belief is justified is determined by whether it is produced by a truth-conducive process (or, more strictly, a process that is conducive to a high truth-ratio). We know from Goldman (1999) that Goldman regards true belief as the fundamental epistemic value, and he is clear that this doesn't mean true belief considered abstractly (e.g., the *property* of being true and a belief), but rather the *state* of true belief. Hence, it would seem to follow that Goldman's view claims that justification is determined by conducivity to a fundamentally epistemically valuable state. One might hence insist on behalf of Berker, Ahlstrom-Vij, and Dunn that whatever Goldman and other reliabilists have said about their own views, these views do count as reliabilist according to their definitions.

But there are three problems with this reasoning. Firstly, it does not follow even from Goldman's simple statement of his view that it is a version of consequentialism according to the aforequoted glosses by Berker, Alstrom-Vij and Dunn. Secondly, as we've already seen, the quoted statement by Goldman is an oversimplification, and the details matter: it is not consequentialist in its more detailed statements. Thirdly, the glosses on epistemic consequentialism are too simple and need expansion to secure an analogy with the consequentialist tradition in ethics. Once expanded, however, it becomes clear in a further way that neither Goldman's view

nor other views classified as consequentialist by Berker are genuinely analogous to consequentialism in ethics. I will close this section by expanding a bit on these three points.

Consequentialism seeks to explain some normative properties and facts (most centrally, rightness and wrongness facts) in terms evaluative properties and facts (most centrally, goodness and badness facts). It is essential for the *explanans* to involve evaluative categories. Yet although true belief has the property of being good, it is not itself a normative property according to most epistemologists. One does not explain the property of epistemic justification in terms of the property of epistemic goodness just by explaining it in terms of some relation to true belief, even an instrumental one. Indeed, if epistemic justification is immediately grounded in some instrumental relation to truth, where the grounding basis includes no normative elements, then it is not true that one normative category is being explained in terms of another. Instead, a normative category is being explained in terms of a *non-normative* category.

In this respect, Goldman (1979: 1)'s own analogy with utilitarianism is slightly misleading.[27] Utilitarianism is standardly understood as a version of consequentialism. But it is inconsistent with non-consequentialist views like Kantianism only if it at least implicitly takes a two-part form, where a consequentialist account of rightness in terms of promotion of the good is first adopted, and then goodness is explained in terms of utility (which might itself be normative if utility = well-being). Indeed, there are versions of utilitarianism—e.g., Hare (1981)'s and Parfit (2011)'s rule utilitarianism—that are derived from more fundamental non-consequentialist foundations. In the case of Hare (1981) these foundations are Kantian, and Parfit's Kantian argument for rule utilitarianism suggests the same could hold for him. A utilitarian view that rejected a normative explanation of rightness in terms of goodness and directly appealed to something clearly non-normative (e.g., happiness) would hence be consistent with some non-consequentialist views.

Similar remarks apply even more forcefully to Berker (2013a: n.27)'s claim that Greco and Sosa circa *Knowledge in Perspective* are epistemic consequentialists. Both theorists *distally* explain epistemic justification and knowledge by appealing to truth-conducive cognitive abilities, which are straightforwardly naturalistic. But they are virtue epistemologists. Sosa has from

[27] It proved misleading to Berker (2013a) in a different way, as Goldman (2015: 136) notes. It was only intended as a formal analogy. He wasn't saying that reliabilism is substantively like utilitarianism.

the beginning emphasized an analogy with Aristotelian virtue ethics. Yet virtue ethics is one of the central alternatives to consequentialism. Part of what makes it an alternative is that it first explains rightness and other normative categories in terms of virtue. There is then disagreement among virtue ethicists about the nature of virtue itself, with Aristotelians favoring a skill-based conception. If virtue epistemology is to parallel virtue ethics, it will also ground justification in virtue and only then give a certain account of virtue, with Sosa and Greco favoring an austere skill-based conception, and responsibilists favoring a richer conception. Hence, the mere fact that Sosa and Greco ultimately want to explain justification and knowledge by appealing to truth-conducive abilities doesn't make them consequentialists, and their self-identification should make one doubt the aptness of this description.

Even if it were right to claim that Goldman's simple description of his view is consequentialist, it wouldn't follow that Goldman's considered view is consequentialist. As I emphasized earlier, it is not. It embraces the epistemic analogue of side-constraints on the promotion of the good in more than one place. Strikingly, Ahlstrom-Vij and Dunn (forthcoming: §3) recognize this fact, but go on to insist that because there are other characterizations of consequentialism that don't mention side-constraints, it follows that Goldman's view might still be consequentialist. While it is true that there are statements of the consequentialism/non-consequentialism divide that don't *explicitly* mention the notion of a side-constraint, it is standard in ethics to assume that embracing side-constraints is sufficient for being a non-consequentialist (though there may be other ways). I hence don't think it is right to call Goldman's view consequentialist in a sense that mirrors ethics.

Ahlstrom-Vij and Dunn (forthcoming: §4) argue that a characterization of consequentialism that they derive from remarks from Smith (2009), Parfit (2011), and Sinnott-Armstrong (2011) seems consistent with embracing side-constraints. This characterization is their 'Dependency' gloss, according to which '[a] theory of rightness is consequentialist iff it maintains that facts about rightness depend solely on the goodness of consequences, be it of individual actions, of accepting or complying with certain rules, or the like' (11). They then observe that theories that restrict the dependency to dependency on immediate consequences can reject problematic trade-offs, and conclude that consequentialism is consistent with side-constraints. But this is a

*reductio* of that gloss on being a consequentialist, and the ethical theorists they mention would agree that a view that essentially incorporates side-constraints couldn't be consequentialist. The reaction to this point should be that the Dependency criterion needs supplementation by the totalizing requirement, which was doubtless in the background of some of these theorists' remarks. But embracing that requirement excludes restriction to immediate consequences, and for good reason: otherwise a key motivation for consequentialism—the irrationality of preferring acknowledged lesser good in what one is directly evaluating—is lost. While one might seek to respect this requirement by tinkering with the value theory and insisting that distal consequences always have neutral value, the result would not be a plausible value theory. Patently bad distal consequences are bad!

# 4 Many Non-Consequentialist Bases for Reliabilism

So far I've given a historical argument for (C1). I want now to abstract away from history and describe four broader ways to give reliabilism non-consequentialist foundations. I'll then explain in §5 why reliabilists should seek some such foundations.

## 4.1 Knowledge (Sort of) First

The first strategy makes explicit what seemed implicit in reliabilism's early history. It proposes that reliability is important in virtue of being the key to the analysis of knowledge. While this strategy puts knowledge axiologically first, it allows that knowledge is analyzable, perhaps adopting a simple analysis along the following lines:

> S knows that p $=_{\text{df-metaphysics}}$ (i) S has a true belief that p, (ii) this belief was formed by a reliable process, and (iii) the very accuracy of the belief manifests the reliability of the process.

Although it takes reliability to be axiologically posterior to knowledge, this view counts as a version of reliabilism in virtue of treating knowledge as metaphysically explained by reliability. This combination of metaphysical priority and axiological posteriority is familiar and perfectly coherent. By analogy, consider how dabs of paint in a good painting get value by constituting it, so that even if they seemed worthless in isolation, they derive value from their role in a whole.

This view could be defended in a non-normative or a normative form. A non-normative version would view knowledge as simply a determinable factive mental state, itself no more normative than more determinate factive mental states like seeing that p and remembering that p. The reliabilist analysis offered by this kind of view would be in the same ballpark as (e.g.) causal theories of intentional action. This view is consistent, however, with taking knowledge to be something fundamentally epistemically good, from which reliability derives value as a key constituent. After all, some candidates for intrinsic value are not normatively constituted. Consider pleasure. Just as whatever constitutes pleasure will be valuable in virtue of constituting something intrinsically valuable, so will whatever constitutes knowledge be valuable in virtue of constituting something fundamentally epistemically valuable.

If one wants to explain the value of justified belief, one could add another idea familiar from knowledge-first epistemology: namely, the idea that the fundamental epistemic norm is the knowledge norm. This norm can itself be understood in a deontology-friendly way, as the following correctness norm (where 'correct' means *objectively permissible*):

**KNC**: It is correct to believe that p if and only if you know that p.

Justification could then be viewed not as a means to true belief but rather as a status had when one is in apparent conformity from one's own perspective with (KNC). Alternatively, if one is willing to deny that there are justified false beliefs, one could view justified belief as coinciding for first-order normative reasons with knowledgeable belief. This isn't to say that justification is a constituent of knowledge, but rather that knowledge is itself a justification-making characteristic analogous in role to the right-making characteristics one finds in normative ethics.

## 4.2 Truth Deontologically First

A second option now becomes apparent. One could put truth rather than knowledge first deontologically, and derive the significance of reliability from the fact that it is the key constituent of *compliance* rather than mere conformity to the truth norm:

**TNC**: It is correct to believe that p if and only if p is the case.

One wants not merely to conform to norms, but to comply with them. The rationale might be the non-consequentialist one that the law calls for one to act *from* it, not merely *in accordance with* it. This is a stronger relation than conformity, which can be accidental. This rationale figures centrally in Herman (1981) and Baron (1995)'s interpretations of Kant on moral worth, and isn't foreign to the non-consequentialist tradition. Provided one offers a psychologically undemanding account of acting 'from' a norm—a long-standing goal of the Kantian literature on moral worth—one can derive accounts of justified belief and knowledge that are extensionally equivalent to familiar reliabilist accounts, but with non-consequentialist foundations. Knowledge could be viewed as consisting in conformity with the truth norm as a manifestation of *respect* for the truth norm, where such respect could be a psychologically undemanding state that Huck Finn has for Jim's humanity. This account would make doubly reliabilist predictions, firstly in virtue of the way it understands respect, and secondly in virtue of the fact that the manifestation relation required for knowledge implies non-accidentality.

## 4.3 Virtue First

A third option is to take the analogy with virtue ethics as an alternative to consequentialism seriously while retaining externalist themes, defending a kind of synthesis of responsibilism and Aristotelian skill-based virtue theory. Zagzebski (1996) came close to offering such a theory. While some aspects of her view are worth dropping, an attractive view can be constructed on the foundations to which she drew attention.

Zagzebski (1996: §§4.1.2-4.1.3) distanced herself from an earlier internalist virtue responsibilist—viz., Montmarquet (1986)—by arguing that virtue has a 'success component'. The rationale is straightforward: virtues are *excellences*, and one cannot have an excellence to $\phi$ unless one tends to succeed in $\phi$-ing when in suitable environments and in suitable shape. Not just any kind of reliable belief-forming disposition merits the title of 'excellence', however. Accordingly, Zagzebski also imposes a motivational requirement on knowledge. Although it is familiar to object to this account on the grounds that it overintellectualizes—well, overpsychologizes—justified belief and knowledge,[28] Zagzebski observed that the motivational constraint

[28] See, e.g., Baehr (2011).

could be understood in a sufficiently undemanding way that the account could turn out extensionally equivalent to sophisticated reliabilist views:

> [M]y definition of knowledge in terms of acts of intellectual virtue can be interpreted in a more externalist way than I have intended, more like Plantinga. This could be done by modifying the motivational element in my account of virtue, making it weaker and farther removed from conscious awareness and control, although I do not think the internalist aspect can be eliminated entirely. The resulting notion of virtue might look a lot like proper function. [...] I mention this, not because I think it is a good idea to bend it [in this way], but because I want to indicate that what I propose here can be adapted by those whose intuitions about knowledge are... more strongly externalist... than mine.[29]

Precisely such a bending of the view was, in effect, wrought by Greco (1999)'s agent reliabilism and Sosa (2015)'s updated reliabilist virtue epistemology, which were intended to reconcile responsibilist intuitions with externalist modesty about the psychology of knowing. Provided that one construes the biconditionals offered by agent reliabilists factorally rather than foundationally, one can then defend these biconditionals on the basis of Zagzebski's foundations together with psychological modesty about appropriate motivation. The advantage that this modified Zagzebskian approach would have over Greco/Sosa-style virtue epistemology would then be fine-grained: given the non-consequentialist foundations for its extensional predictions, it has a better solution to reliabilism's value problem.

## 4.4 Objective Reasons First

The final approach I'll mention is an updated version of indicator reliabilism that profits from insights in recent work on reasons and rationality outside epistemology.[30] In ethics, it has become common to view the fundamental notion of a reason as that of a fact, often worldly, that counts in favor of some action or attitude. On one attractive view, justification of the sort that contrasts with excuse is then taken to be a function of the reasons that one possesses.

The relation of possession can, some have suggested,[31] be analyzed in terms of a non-normative access condition together with a normative condition of competence with objective reasons. Since competence with objective reasons plausibly coincides extensionally with sensi-

[29] Zagzebski (1996: 329-330).

[30] This approach has precedent: McDowell (1994) and Fricker (2007) developed such an account of the epistemology of testimony, and generalizing their accounts suggests the view described below.

[31] See Sylvan (2016) and Sylvan (Forthcoming).

tivity to reliable indicators—understood to include worldly indicators, not just the mental states of classic indicator reliabilism—and access can be understood in a non-normative reliabilist fashion, this account will be extensionally equivalent to a sophisticated indicator reliabilism.

But the foundations of this account needn't be consequentialist. On the one hand, one could take objective reasons as primitive and see the truth that objective epistemic reasons for belief are truth-indicators as a synthetic *a priori* truth rather than a clue to their analysis. On the other hand, one could analyze objective reasons in any number of patently non-consequentialist ways that are consistent with the account's remaining an indicator reliabilism. For example, one could hold that objective reasons to $\phi$ in a domain are probability-raisers of the proposition that $\phi$-ing would be correct (in a normative sense) in the domain. And one could then understand the norms of correctness in a domain deontologically.

While this account appeals to the probability that $\phi$-ing is correct, it doesn't follow that it is a consequentialist account. Every complete normative theory must appeal to some such probability somewhere, since every normative theorist should recognize a distinction between objective correctness in $\phi$-ing and $\phi$-ing that is justified, praiseworthy, etc., where there is a probabilistic relation between the two. For example, Kantians can agree that there is a distinction between objective rightness and justification, viewing objective rightness as action in compliance with the Categorical Imperative and justified action as action that is likely relative to the indicators within one's perspective to be in compliance with the Categorical Imperative. The core distinction is a structural one neutral on substantive normative foundations.

## 4.5 Do These Accounts Really Count as Reliabilist?

We discover an embarrassment of riches if we seek non-consequentialist foundations for accounts that extensionally coincide with various reliabilisms. But one might still ask whether these accounts are reliabilist, perhaps with the hope of resisting my proof by illustration that reliabilism can have non-consequentialist foundations. While the raw materials for answering this question were provided earlier, let's walk through why these accounts merit the title of 'reliabilism'.

A broader point is in order first. Several of the foregoing accounts were framed as 'X-first'

accounts, where X $\neq$ reliability. It doesn't follow from this mere fact that these accounts are not reliabilist. Reliabilists needn't think that reliability is absolutely first. In the weak, non-Williamsonian sense in which the first proposal was 'knowledge-first', the original versions of reliabilism were knowledge-first. Similarly, in the sense in which the third proposal was 'virtue-first', Greco's agent reliabilism and Sosa's new version of virtue reliabilism are virtue-first. And none of the alleged examples of consequentialist reliabilism are reliability-first views either. On these views, *the value of accuracy* comes first. It is only when we consider how epistemic norms derive from the value of truth that we encounter reliability.

Let's now wade more carefully through the reasons why the foregoing accounts are genuine reliabilisms. As I noted at the outset, reliabilism can be offered as a non-normative theory or a normative theory, and if it is offered as normative theory, it can be offered factorally or foundationally. We can take each possibility in turn.

A non-normative version of reliabilism would propose that a certain kind of reliable ancestry is the key element in the constitution of knowledge, itself understood as no more normative than perception. If this view deserves the name 'reliabilism'—as the history suggests it does—then the first proposal deserves the name 'reliabilism'.

If a factoral theory can deserve the name 'reliabilism', then the second through fifth accounts all straightforwardly deserve the name of 'reliabilism' as well. Factoral theories are consistent with multiple foundations. In virtue of predicting that the same factors matter to justification as uncontroversially reliabilist theories, the second through fifth accounts count as factorally reliabilist.

It is only if reliabilism is understood exclusively as a foundational normative theory that the foregoing accounts might not merit the name. But this exclusive criterion prevents *classic reliabilist views* from counting as genuine reliabilisms. For some classic reliabilist views weren't intended as foundational normative theories or emphasized different normative notions (e.g., virtue) at the foundational normative level. And again, while Goldman (1986) did offer reliabilism as a foundational normative theory akin to consequentialism, his view contained restrictions that are ill-motivated from a consequentialist perspective. A generalization of this last point indicates why reliabilists ought to be non-consequentialists, as we will now see.

# 5 Why Reliabilists Ought to Be Non-Consequentialists

## 5.1 The Argument from Elimination

Here in outline is an argument that reliabilists shouldn't be consequentialists:

*The Argument from Elimination*

1. Reliabilism is a kind of consequentialism only if it is a foundational normative theory.
2. But reliabilism should not be defended as a foundational normative theory: if it is defensible, it is only as a factoral theory or a non-normative theory.
3. So, reliabilism shouldn't be defended as a kind of epistemic consequentialism.
4. Moreover, epistemic consequentialism is inconsistent with the most plausible factoral and non-normative versions of reliabilism.
5. So, not only should reliabilism not be offered as a version of consequentialism, it should be offered in opposition to epistemic consequentialism.

I already defended (1) in §2 when I explained why consequentialism should be classified as a foundational normative theory. But here is a further case for (1). Reliabilism can be offered in three formats. The factoral and non-normative formats are consistent with epistemic non-consequentialism. To be sure, to say that a version of reliabilism is consistent with non-consequentialism isn't to say that it is inconsistent with consequentialism in its most plausible formulations, as (4) claims. But if a theory T is consistent with both the affirmation and negation of another theory T*, T cannot be a *version* of T*. Hence, (1) is true.

(2) and (4) need defense. I previewed the defense of (4) in §3 when I noted that the restrictions that careful reliabilists have wanted to impose are not restrictions any honest-to-goodness epistemic consequentialism—i.e., one analogous to consequentialism in ethics—would contain. I will deepen this argument by abstracting away from the details of the particular reliabilist accounts I discussed there, arguing that this point holds for any plausible version of reliabilism.

I have not previewed the defense of (2), which is an importantly different claim from (4). (2) doesn't claim that reliabilism is implausible when framed as a form of epistemic consequentialism. It rather claims that reliabilism—setting the question of whether it is a kind of consequentialism aside—is implausible as a foundational theory. But the case for this claim is not unrelated to the case for (2): a foundational theory that puts reliability first would have

the same difficulties motivating restrictions needed for factoral plausibility as epistemic consequentialism.

## 5.2 In Defense of (2)

Let's consider what a foundational first-order reliabilism would look like. Like consequentialism, this view could take direct and indirect forms. A direct first-order foundational reliabilism about a certain epistemic status S (e.g., justification) of a certain potential target of evaluation T (e.g., a doxastic attitude) in a world *w* would make the following schematic claim:

> **Schematic Direct Foundational Reliabilism**: The fundamental normative explanation of whether target T has status S in *w* is that T is reliable in *w*.

Where the relevant targets of evaluation are doxastic attitudes, no prominent reliabilists have favored such a view. This view doesn't claim that it is the reliability of the belief-forming process or method that explains why the beliefs are justified, but rather the reliability of the *beliefs* formed by this process (whatever that might mean), independently of the method by which they are formed.

Perhaps the closest relative of such a view would be a simple counterfactual reliabilism, since in its simple form counterfactual reliabilism does ask us to consider the behavior of the belief in various counterfactual scenarios. But simple views of this kind face obvious counterexamples, like Nozick's (1981)'s grandmother case and Goldman (1976, 1983)'s dachshund case, as well as in-principle problems like closure failure of the most elementary sort (see Kripke (2011: Ch.7)). The former problems are why Nozick relativized to methods, and the latter why Roush (2005) has abandoned the Nozickian account for non-basic knowledge and offered a recursive account which is Nozickian only at the foundational level.

It is no surprise that reliabilists about justified belief and knowledge do not appeal to the reliability of beliefs but rather the reliability of the processes, methods, dispositions, etc., which generate beliefs. So if they were to commit to foundational reliabilism, it would have to take the following different form:

> **Schematic Indirect Foundational Reliabilism**: The fundamental normative explanation of whether a target T has status S in *w* is that T* is reliable in *w*, where T stands in an

indirect relation R to T* (e.g., T is produced by T*).

While there are extensional concerns with this theory, the deeper concerns center around its claim of fundamental normative explanation.

My core concern is hardly without precedent, though it hasn't been put in quite the way I'll put it. As others have noted,[32] the way that some prominent reliabilists address objections like BonJour (1980)'s clairvoyance objection and defeat objections seem ill-motivated if reliability is what matters fundamentally.[33] Some of these writers accordingly suggest that we ditch reliabilism for a different kind of theory. But such ditching is advisable only if reliabilism is a foundational theory, something which, as we've seen, it needn't be. The refined theories prominent reliabilists have offered are defensible as factoral theories. Since externalists have tended to put a premium on commonsense intuition, often contenting themselves with a description of the epistemic concepts revealed in our ordinary practice of attribution,[34] it is hard to see why they wouldn't be willing to rest content with limiting their ambitions to the factoral level. So, it is challenging to be a foundational reliabilist if one makes consistency with ordinary intuition a desideratum, though factoral reliabilism of the sort defended by central reliabilists is not so challenged.

Let's consider the concern in a bit more detail. Two most familiar problems for simple forms of reliabilism about justified belief are the clairvoyance and new evil demon problems. The former can be addressed extensionally by requiring not just real reliability but apparent reliability from one's perspective. The latter can be addressed extensionally in two non-revisionary ways: by indexing reliability to the actual world, or by allowing that apparent reliability from one's perspective suffices for some kind of normative status, though perhaps weaker than full-fledged justification (cf. Goldman 1988). If we prefer foundational reliabilism, it is unclear why one would favor either of these moves. If reliability provides the most fundamental expla-

[32]See, e.g., Greco (2010: 158) and Beddor (2015).

[33]Bergmann (1997) nicely documents this strand in the literature. But it has seen further iterations. Goldman (2011), for example, allows that mere experiences are defeaters even when they are misleading (and indeed, at least conceivably, not reliable indicators at all), an allowance that parallels his (1986) allowance of mere beliefs as defeaters (which is also allowed by Plantinga (1993) and Nozick (1981)).

[34]The Moorean tendency among many externalists to reject out of hand theories that imply skepticism is an important illustration. While Goldman (1993) envisaged two projects—descriptive epistemology, which seeks to understand our 'epistemic folkways', and normative epistemology, which seeks to refine them—his later work on methodology (e.g., Goldman and Pust (1998) and Goldman (2007)) suggests that he takes philosophy's proper task to be (more or less) a description of our ordinary concepts as revealed by our practices of attribution.

nation of why a belief has normative standing, apparent reliability ought to make for apparent justification, not any real kind of justification. Similarly, the fact that a process is unreliable in the believer's world should prevent it from being justified in that world, whatever might be said of the process in other worlds. Finally, if a process is reliable in the believer's world, it is unclear why it also must be apparently reliable from the believer's perspective to qualify for full-fledged normative status: everything needed for a foundational explanation of its normative status is already in place. One could revise the story, holding that the complete explanation of how a belief acquires its normative status must advert to apparent reliability. But that story is implausible as a foundational story. If apparent reliability has significance, its significance is derivative, not foundational.

## 5.3 In Defense of (4)

The argument for (4) turns on similar points, though the problems for giving reliabilism consequentialist foundations are more pronounced than the problems for a foundational reliabilism that doesn't claim the 'consequentialist' moniker. For while reliabilism is the property of epistemologists and can perhaps be framed however its leading defenders fancy, consequentialism is not. Given consequentialism's roots in ethics, an epistemological theory worthy of the name 'consequentialism' should parallel theories by that name in ethics. Otherwise 'consequentialism' threatens to become a mere homonym in ethics and epistemology; even if epistemologists aren't bugged by this threat, it is unclear why we should bother introducing a misleading new label for a theory that already has a name. But the stronger one makes the analogy between consequentialism in ethics and consequentialism in epistemology, the farther away ones move from a theory reliabilists will gladly endorse.

Some of the problems here have already been well-documented in the literature attacking epistemic consequentialism.[35] I'm not sure we should be persuaded by Selim Berker's much-discussed arguments if understood as arguments against familiar views in epistemology. But Berker does persuasively show that theories of epistemic justification that closely parallel consequentialism in ethics will at best be highly revisionary, revisionary enough to

[35] See Berker (2013a-b), Fumerton (2001) and Firth (1981).

fail the desideratum of continuity with ordinary attribution that Goldman (1993) imposes on 'scientific' epistemology. This combination of thoughts recommends my conclusion. While epistemic consequentialism is probably false, reliabilism remains unscathed because it never plausibly was consequentialist in the first place.

Even if reliabilism had been loudly promulgated as a consequentialist view in some formulations, it shouldn't have been: a firm analogy with consequentialism in ethics won't secure the predictions reliabilists have wanted or motivate the restrictions they have imposed. Ahlstrom and Dunn (2014)'s response to Berker provides a nice illustration of this point, as I noted earlier. They are right that Goldman's process reliabilism isn't impugned by Berker's arguments. But again, these restrictions have no parallel in consequentialist theory in ethics, and look more like the agent-centered restrictions of deontology. No principled consequentialist would hold the disvalue of violating a restriction to be so great that the value of a certain set of consequences could never outweigh it. Yet only such a claim makes consequentialist sense of Goldman's restrictions (which, while plausible, is not sustainable on genuinely consequentialist foundations).

But we needn't invoke Berker-style arguments to defend (4). We need only consider where epistemic consequentialist foundations will lead the reliabilist *vis-à-vis* clairvoyants, demon worlders, and the value of justified and knowledgeable belief. While veritist epistemic consequentialists might exploit the seemingly theory-neutral distinction between subjective and objective rightness to address some intuitions about clairvoyance cases and the demon world, there are others they will be ill-equipped to address. At best they will be able to claim that the clairvoyant believes in a way that is subjectively wrong though objectively right, and that the demon worlder believes in a way that is subjectively right through objectively wrong. But they will not have the resources to explain why we think the demon-worlder is *genuinely better off from an epistemic point of view* in virtue of believing rationally.

While subjective rightness may be a kind of rightness, expected value is not a kind of value. To a veritist epistemic consequentialist, the demon-worlder's rational worldly belief-forming processes should seem bad all-things-considered, since they only reliably produce false beliefs. A related claim holds vis-à-vis the clairvoyant: to a veritist epistemic consequentialist,

there shouldn't seem to be anything bad about the clairvoyant's belief-forming dispositions, since they reliably produce true beliefs. Even if the veritist epistemic consequentialist could find something damning to say about the clairvoyant's belief-forming processes and something good to say about the demon-worlder's belief-forming processes (in her world), the core of the swamping problem would remain vis-à-vis our intuitions about the value of the *products* of these processes. For as Jones (1997) noted in giving the swamping problem one of its earliest formulations, the problem is a general problem for epistemic instrumentalism and other theories founded on its basis, not a problem for reliabilism as such. Indeed, if reliabilism is offered factorally and given virtue-theoretic or deontological foundations, it won't face the problem at all.

Now, there is a familiar move that consequentialists in ethics make to address intuitive deficiencies of this kind: they expand their theory of intrinsic value, moving from a monistic axiology like hedonism to a pluralist axiology like Moore's. But the analogue of this move in epistemology threatens to undermine the reliabilist's explanatory ambitions vis-à-vis the nature of justification, rationality, and knowledge. It is hence not a move that a reliabilist would be well-advised to make. Suppose, for example, we expand our account of intrinsic epistemic value to include rational belief, justified belief, and knowledge. If we do so and still seek to derive reliabilism from epistemic consequentialism, we will end up with a correspondingly expanded conception of reliability. Reliability won't simply be truth-conducivity but conductivity to a plurality of intrinsic epistemic values. If we are serious about deriving justification from consequentialist foundations, our account of justified belief will have to be correspondingly expanded. A justified belief by pluralist epistemic consequentialist lights will be one that is formed by a process that is reliable in the expanded sense. The result is that we are left with a circular account of the properties reliabilism promised to analyze.

I conclude, then, that if reliabilists want to avoid implausibility, they will need to seek non-consequentialist foundations. Hence, not only should reliabilism not be offered as a version of epistemic consequentialism, but it should also be founded positively on a kind of non-consequentialism.

# 6 Concluding Remarks

Of course, there remains a lot more work to do in developing a non-consequentialist reliabilism, but the goal of this paper has just been to explain why epistemologists attracted to reliabilism should head in this direction. Let's take stock of how this goal has been achieved.

I pursued this goal in two stages, beginning with the simpler task of explaining why reliabilism's fate needn't depend on epistemic consequentialism's. In §2, I laid the foundations for the divorce of reliabilism and epistemic consequentialism by drawing attention to a distinction between three kinds of theorizing: non-normative theorizing about the metaphysical constitution of epistemic properties, on the one hand, and normative theorizing at the factoral and foundational levels, on the other hand. Epistemic consequentialism is an example of the last kind of theorizing, and reliabilism would be independent of epistemic consequentialism if it were understood as an example of either of the other two kinds of theorizing. I then argued in §3 that as a matter of historical fact, reliabilism has not primarily been a foundational normative theory. It started out as a piece of non-normative theorizing about the metaphysics of knowledge, and when it took a normative turn at the end of the 1970s, it wasn't immediately under the rubric of foundational normative theorizing. It was only in one admittedly important contribution to reliabilism in the 1980s that reliabilism took a foundational normative form allegedly similar to consequentialism. But even this form turned out to be disanalogous to the familiar versions of consequentialism.

Of course, one might think that the historical record is a record of failure, and that the fact that reliabilism often wasn't a version of epistemic consequentialism says little about the theoretical format in which it should be couched. But as I noted in §4, there are many possible non-consequentialist foundations for reliabilism, and some of them are more attractive than the supposedly consequentialist foundations offered in Goldman (1986). Furthermore, there is a deeper argument, as we saw in §5, for restricting one's attention to these kinds of options. Reliabilism *shouldn't* be offered as a version of consequentialism, since the only version of reliabilism that would coincide with a recognizably consequentialist view would be a foundational first-order version which fails to provide principled foundations for the factors that commonsense intuition suggest to be important. Epistemic consequentialism is a poor basis for

the most plausible versions of reliabilism. Hence, reliabilists ought to work in opposition to epistemic consequentialists. To do so, they will need to seek a partnership with some foundational non-consequentialist outlook. Here there is an embarrassment of riches, as I had noted in §4, though it would take another paper to identify the best option.

# 7 References


Ahlstrom-Vij, K. and Dunn, J. 2014. 'A Defence of Epistemic Consequentialism.' *Philosophical Quarterly* 64: 541-551.

Ahlstrom-Vij, K. and Dunn, J. 2018. 'Introduction' to *Epistemic Consequentialism*. Oxford University Press.

Ahlstrom-Vij, K. and Dunn, J. Forthcoming. 'Is Reliabilism a Form of Epistemic Consequentialism?' *American Philosophical Quarterly*.

Alston, W. P. 1985. 'Concepts of Epistemic Justification.' *The Monist* 68: 57-89.

Alston, W. P. 1988. 'An Internalist Externalism.' *Synthese* 74: 265-83.

Armstrong, D. M. 1961. *Perception and the Physical World*. London: Routledge.

Armstrong, D. M. 1968. *A Materialist Theory of the Mind*. London: Routledge.

Armstrong, D. M. 1973. *Belief, Truth and Knowledge*. Cambridge University Press.

Baehr, J. 2011. *The Inquiring Mind*. Oxford University Press.

Bales, R. E. 1971. 'Act-Utilitarianism: Account of Right-Making Characteristics or Decision-Making Procedure?' *American Philosophical Quarterly* 8: 257–265.

Baron, M. 1995. *Kantian Ethics Almost without Apology*. Ithaca: Cornell University Press.

Beddor, R. 2015. 'Process Reliabilism's Troubles with Defeat.' *Philosophical Quarterly* 65: 145-159.

Benson, P. 1987. 'Moral Worth.' *Philosophical Studies* 51: 365-382.

Bergmann, M. 1997. 'Internalism, Externalism, and the No-Defeater Condition.' *Synthese* 110: 399-417.

Berker, S. 2013a. 'Epistemic Teleology and the Separateness of Propositions.' *Philosophical Review* 122: 337-393.

Berker, S. 2013b. 'The Rejection of Epistemic Consequentialism.' *Philosophical Issues* 23: 361-387.

Berker, S. Forthcoming. 'The Unity of Grounding.' *Mind*.

BonJour, L. 1980. 'Externalist Theories of Empirical Knowledge.' *Midwest Studies in Philosophy* 5: 53-73.

Copp, D. 2009. 'Toward a Pluralist and Teleological Theory of Normativity.' *Philosophical Issues* 19: 21-37.

Copp, D. 2014. 'Indirect Epistemic Teleology Explained and Defended' in Fairweather, A. and Flanagan, O. (eds.) *Naturalizing Epistemic Virtue*. Cambridge University Press.

Dean, R. 2006. *The Value of Humanity in Kant's Moral Theory*. Oxford University Press.

Dreier, J. 1993. 'Structures of Normative Theories.' *The Monist* 76: 22-40.

Dretske, F. 1971. 'Conclusive Reasons.' *Australasian Journal of Philosophy* 49: 1-22.

Dretske, F. 1981. *Knowledge and the Flow of Information*. Cambridge, MA: The MIT Press.

Dunn, J. 2015. 'Epistemic Consequentialism.' *Internet Encyclopedia of Philosophy*. ISSN 2161-0002. URL = <https://www.iep.utm.edu/epis-con/>.

Firth, R. 1981. 'Epistemic Merit, Intrinsic and Instrumental.' *Proceedings and Addresses of the American Philosophical Association* 55: 5-23.

Foley, R. 1987. *The Theory of Epistemic Rationality*. Cambridge, MA: Harvard University Press.

Foot, P. 1985. 'Utilitarianism and the Virtues.' *Mind* 94: 196–209.

Fricker, M. 2007. *Epistemic Injustice*. Oxford University Press.
Fumerton, R. 2001. 'Epistemic Justification and Normativity' in Steup, M. (ed.) *Knowledge, Truth, and Duty*. Oxford University Press.
Goldman, A. I. 1967. 'A Causal Theory of Knowing.' *Journal of Philosophy* 64: 357-372.
Goldman, A. I. 1975. 'Innate Knowledge' in Stich, S. (ed.) *Innate Ideas*. University of California Press.
Goldman, A. I. 1976. 'Discrimination and Perceptual Knowledge.' *Journal of Philosophy* 73: 771-791.
Goldman, A. I. 1979. 'What Is Justified Belief?' in Pappas, G. (ed.) *Justification and Knowledge*. Dordrecht: D. Reidel.
Goldman, A. I. 1983. 'Review of Philosophical Explanations.' *Philosophical Review* 92(1): 81-88. Goldman, A. I. 1986. *Epistemology and Cognition*. Cambridge, MA: Harvard University Press.
Goldman, A. I. 1988. 'Strong and Weak Justification.' *Philosophical Perspectives* 2: 51-69.
Goldman, A. I. 1993. 'Epistemic Folkways and Scientific Epistemology.' *Philosophical Issues* 3: 271-285. Goldman, A. I. 1999. *Knowledge in a Social World*. Oxford University Press.
Goldman, A. I. 2007. 'Philosophical Intuitions: Their Target, Their Source, and Their Epistemic Status.' *Grazer Philosophische Studien* 74: 1-26.
Goldman, A. I. 2011. 'Toward a Synthesis of Evidentialism and Reliabilism?' in Dougherty, T. (ed.) *Evidentialism and its Discontents*. Oxford University Press.
Goldman, A. I. 2012. *Reliabilism and Contemporary Epistemology*. Oxford University Press.
Goldman, A. I. and Beddor, R. 2015. 'Reliabilist Epistemology' in Zalta, E. (ed.) *The Stanford Encyclopedia of Philosophy (Winter 2015 Edition)*. URL = <http://plato.stanford.edu/ archives/win2015/entries/reliabilism>
Goldman, A. I. and Pust, J. 1998. 'Philosophical Theory and Intuitional Evidence' in Depaul, M. and Ramsey, W. (eds.) *Rethinking Intuition*. Lanham: Rowman & Littlefield.
Greco, J. 1999. 'Agent Reliabilism.' *Philosophical Perspectives* 19: 273-296.
Greco, J. 2010. *Achieving Knowledge*. Cambridge: Cambridge University Press.
Hare, R. M. 1981. *Moral Thinking*. Oxford University Press..
Herman, B. 1981. 'On the Value of Acting from the Motive of Duty.' *Philosophical Review* 66: 359-382.
Herman, B. 1993. *The Practice of Moral Judgment*. Cambridge, MA: Harvard University Press.
Herman, B. 2007. *Moral Literacy*. Cambridge, MA: Harvard University Press.
Jones, W. 1997. 'Why Do We Value Knowledge?' *American Philosophical Quarterly* 34: 423-439.
Kagan, S. 1992. 'The Structure of Normative Ethics.' *Philosophical Perspectives* 6: 233-242.
Kagan, S. 1997. *Normative Ethics*. Westview Press.
Kamm, F. 1992. 'Non-Consequentialism, the Person as an End-in-Itself, and the Significance of Status.' *Philosophy and Public Affairs* 21(4): 354-389.
Kant, I. 1785/2012. *Groundwork for the Metaphysics of Morals*. Cambridge University Press.
Korsgaard, C. 1996. *Creating the Kingdom of Ends*. Cambridge University Press.
Kripke, S. 2011. *Philosophical Troubles*. Oxford: Oxford University Press.
Littlejohn, C. 2014. 'Fake Barns and False Dilemmas.' *Episteme* 11: 369-389.
Lord, E. 2018. *The Importance of Being Rational*. Oxford University Press.
Lyons, J. 2009. *Perception and Basic Beliefs*. Oxford University Press.
Mantel, S. 2018. *Determined by Reasons: A Competence Account of Acting for a Normative Reason*. London: Routledge.
Markovits, J. 2010. 'Acting for the Right Reasons.' *Philosophical Review* 119: 201-242.
McDowell, J. 1994. 'Knowledge by Hearsay' in Matilal, B. K. and Chakrabarti, A. (eds.) *Knowing from Words*. Dordrecht: Kluwer.
Millar, A. 2010. 'Knowledge and Recognition' in *The Nature and Value of Knowledge: Three Investigations*. Oxford University Press.
Montmarquet, J. 1986. 'Epistemic Virtue.' *Mind* 96: 482-97.
Nozick, R. 1981. *Philosophical Explanations*. Cambridge, MA: Belknap Press.
Parfit, D. 2011. *On What Matters*. Oxford University Press.
Pettit, P. 1991. 'Consequentialism' in Singer, P. (ed.) *A Companion to Ethics*. Malden: Blackwell.

Pettit, P. 1997. 'The Consequentialist Perspective' in Baron, M., Pettit, P. and Slode, M. (eds.) *Three Methods of Ethics: A Debate*. Malden: Blackwell.
Price, H. H. 1953. *Thinking and Experience*. London: Hutchinson & Co.
Pritchard, D. 2012. 'Anti-Luck Virtue Epistemology.' *Journal of Philosophy* 109: 247-279.
Railton, P. 1984. 'Alienation, Consequentialism, and the Demands of Morality.' *Philosophy and Public Affairs* 13: 134–171.
Ramsey, F. P. 1931a. 'Knowledge' in Braithwaite, R. B. (ed.) *Foundations of Mathematics and Other Logical Essays*. London: Routledge.
Ramsey, F. P. 1931b. 'Truth and Probability' in Braithwaite, R. B. (ed.) *Foundations of Mathematics and Other Logical Essays*. London: Routledge and Kegan Paul.
Roush, S. 2005. *Tracking Truth*. Oxford: Oxford University Press.
Scanlon, T. M. 1998. *What We Owe to Each Other*. Cambridge, MA: Harvard University Press.
Sellars, W. 1956. 'Empiricism and the Philosophy of Mind' in Feigl, H. and Scriven, M. (eds.) *Minnesota Studies in the Philosophy of Science*. Minneapolis: University of Minnesota Press.
Sidgwick, H. 1907. *Methods of Ethics*. London: Macmillan.
Sinnott-Armstrong, W. 2011. 'Consequentialism' in (ed.) Zalta, E. *Stanford Encyclopedia of Philosophy*. Available at: http://plato.stanford.edu/archives/spr2014/entries/consequentialism/.
Smith, M. 2009. 'Two Kinds of Consequentialism.' *Philosophical Issues* 19: 257–272.
Sosa, E. 1991. *Knowledge in Perspective*. Cambridge: Cambridge University Press.
Sosa, E. 2007. *Apt Belief and Reflective Knowledge, Volume 1*. Oxford: Oxford University Press.
Sosa, E. 2015. *Judgment and Agency*. Oxford: Oxford University Press.
Stratton-Lake, P. 2000. *Kant, Duty, and Moral Worth*. London: Routledge.
Star, D. 2015. *Knowing Better*. Oxford: Oxford University Press.
Swain, M. 1981a. 'Justification and Reliable Belief.' *Philosophical Studies* 40: 389-407.
Swain, M. 1981b. *Reasons and Knowledge*. Ithaca: Cornell University Press.
Swain, M. 1985. 'Justification, Reasons, and Reliability.' *Synthese* 64: 69-92.
Sylvan, K. 2016. 'Epistemic Reasons I: Normativity.' *Philosophy Compass* 11(7): 364–376.
Sylvan, K. 2018. 'Knowledge as a Non-Normative Relation.' *Philosophy and Phenomenological Research* 97: 190-222.
Thomson, J. J. 2008. *Normativity*. Chicago: Open Court.
Unger, P. 1968. 'An Analysis of Factual Knowledge.' *Journal of Philosophy* 65: 157-170.
Vogel, J. 2000. 'Reliabilism Leveled.' *Journal of Philosophy* 97: 602-623.
Wood, A. 1999. *Kant's Ethical Thought*. Cambridge: Cambridge University Press.
Zagzebski, L. 1996. *Virtues of the Mind*. Cambridge: Cambridge University Press.