BIROn - Birkbeck Institutional Research Online

WHY NO TRUE RELIABILIST SHOULD ENDORSE RELIABILISM

Kristoffer Ahlstrom-Vij

Department of Philosophy

Birkbeck, University of London

k.ahlstrom-vij@bbk.ac.uk

Jeffrey S. Dunn

Department of Religion, Philosophy, and History

University of Agder

jeffrey.dunn@uia.no

Abstract: Critics have recently argued that reliabilists face trade-off problems, forcing them to condone intuitively unjustified beliefs when they generate lots of true belief further downstream. What these critics overlook is that reliabilism entails that there are *side-constraints* on belief-formation, on account of which there are some things you should not believe, even if doing so would have very good epistemic consequences. However, we argue that by embracing side-constraints the reliabilist faces a dilemma: she can either hold on to reliabilism, and with it aforementioned side-constraints, but then needs to explain why we should allow the pursuit of justification to get in the way of the acquisition of true belief; or she can deny that there are side-constraints—and in effect give up on reliabilism. We'll suggest that anyone moved by the considerations that likely attract people to reliabilism in the first place—the idea the true belief is good, and as such should be promoted—should go for the second horn, and instead pursue a form of *epistemic utilitarianism*.

1. Reliabilism and the Intrapersonal Trade-off Problem

Consequentialists believe that what's *right* should be understood in terms of what's *good*. For example, for the classic *utilitarian*, it's right to give to charity when it maximizes happiness. Similarly, in epistemology, the *reliabilist* (e.g., Goldman 1979) believes the following:

RELIABILISM: A belief is justified only if it is formed by a process that tends to issue in true belief.[1]

Recently, opponents of reliabilism have suggested that this similarity lands her in trouble. Utilitarians infamously face interpersonal trade-offs where the suffering of some must be traded for the benefit of others (e.g., Thomson 1976). According to her critics, the reliabilist faces *intra*personal trade-offs, where unjustified beliefs must be formed by a person to increase her accuracy with respect to future beliefs (Berker 2013*a*, *b*, Littlejohn 2012, Jenkins 2007, Firth 1981; cf. Greaves 2013).

However, these critics are mistaken. For one thing, the supposed trade-off cases put forward to date do not present a problem for the reliabilist; they're all either not trade-offs the reliabilist needs to make, or not problematic trade-offs (Ahlstrom-Vij and Dunn 2014; Goldman 2015). For example, a simple trade-off problem would consist in the reliabilist needing to condone a belief formed in light of excellent evidence to the contrary, that would have as a causal consequence a great many true beliefs in the future. But since the reliabilist evaluates the justificatory status of beliefs, not with reference to *its* consequences, but with reference to the direct (more on this below) consequences of the type of process that generates it, the reliabilist doesn't have to condone the formation of such a belief, since forming a belief in light of excellent evidence to the contrary arguably constitutes an unreliable process.

---

[1] We have stated this as a necessary condition, since this is something that all reliabilists will agree on and because this is all we need for our arguments to go through. Reliabilists tend to add to RELIABILISM, however, to yield necessary *and sufficient* conditions for justification. For instance, Goldman (1979) accepts this necessary condition and adds a kind of 'no defeater' clause to get a sufficient condition. Similarly, this necessary condition is part of his account of *strong justification* (Goldman 1992*c*); his *normal worlds reliabilism* (Goldman 1986), where justification is a function of reliability in the actual world (assumed to be normal), rigidified for all words; and his *virtue reliabilism* (Goldman 1992*b*), where virtues earn their keep by being reliable.

Further, and perhaps more importantly, while reliabilism is indeed a form of consequentialism, it's not of a kind on which we should even expect trade-off problems to arise in the first place (Dunn and Ahlstrom-Vij 2017). The type of consequentialism on which we should expect trade-off problems is one that doesn't impose any *side-constraints* (e.g., Nozick 1981). In ethics, imposing side-constraints on actions is to maintain that it can be wrong to do something, even if it has very good (including the best) consequences (e.g., because it violates people's rights). In epistemology, imposing side-constraints amounts to holding that there are some things you should not believe, even if doing so would have very good (including the best) epistemic consequences. But the reliabilist accepts that. She holds that if a belief is formed by an unreliable process, then it is not justified. Moreover:

> NORMATIVE GLOSS: If someone's belief that *p* is or would be *unjustified*, then she *should not* believe that *p* (in that way, at that time).[2]

Consider, for instance, the following from Alvin Goldman, arguably the most prominent reliabilist:

> Calling a belief justified implies that it is a *proper* doxastic attitude, one to which the cognizer has an epistemic right or entitlement. These notions have a strong deontic flavor […]. They are naturally captured in the language of 'permission' and 'prohibition' (Goldman 1986: 59).

If Goldman is right—and we suspect most reliabilists would find little to object to here—then RELIABILISM entails:

---

[2] As the quotation from Goldman immediately below this makes clear, the 'should' is to be understood as an *epistemic* sense, not in a more general all-things-considered sense. The same holds for future uses of 'should' in this paper.

SIDE-CONSTRAINTS: There are some things you should not believe, even if doing so

would have very good (including the best) epistemic consequences.

Specifically, if a belief would be formed by way of an unreliable process, that's a belief you should

not hold (in that way, at that time), even if it would have very good epistemic consequences. Hence,

it shouldn't come as a surprise that reliabilism doesn't fall prey to intrapersonal trade-off problems.

That said, we will argue that it's not clear that the reliabilist should *want* to embrace SIDE-

CONSTRAINTS. We will suggest that reflecting on realistic cases in which the best way to

promote true belief in society at large is by having people form unjustified beliefs makes clear that

the reliabilist faces a dilemma. On the one hand, she can hold on to reliabilism, and the prohibition

against unreliably formed but truth-conducive beliefs. But in so doing, she must explain why we

should allow the pursuit of justification to get in the way of the acquisition of true belief, all while

holding that justification is valuable merely as a means to true belief. On the other hand, she can

maintain that it is the acquisition of true belief that matters, but then drop the prohibition against

unreliably formed but truth-conducive belief. In so doing, however, the reliabilist gives up on

reliabilism. We'll suggest that anyone moved by the considerations that likely attract people to

reliabilism in the first place—the idea the true belief is good, and as such should be promoted—

should go for the second horn of that dilemma, and instead pursue a form of *epistemic utilitarianism*.

Before fleshing out the details of our argument, we should make a point about our target.

Much of our attention will be on Goldman, and thereby on what one might think of as *classic*

reliabilism, as captured by RELIABILISM. However, we won't be engaging directly with one

prominent form of reliabilism in contemporary epistemology, namely *virtue reliabilism* (e.g., Sosa

2015 and 2009; Greco 2010). The reason is not that we think that virtue reliabilism is not a form

of reliabilism, but that in previous work we have already argued that virtue reliabilism is reducible

to classic reliabilism.[3] Repeating the relevant arguments here would consume a substantial amount of space that we need in order to achieve the main goal of the paper: to offer a sustained argument against (classic) reliabilism, and provide the outlines of the type of view we believe reliabilists—virtue reliabilists included—should pursue in its stead.

2. Side-constraints and Aggregation

In fleshing out the dilemma for the reliabilist, it's helpful to turn to Goldman's social epistemology. Goldman suggests that '[s]ocial practices are epistemically desirable to the extent that they promote epistemically preferred belief-profiles' (1992*a*: 194). What profiles are *preferred*? 'The central epistemological concepts of appraisal, I argue, invoke *true belief* as their ultimate aim. So the evaluation of epistemic procedures, methods, processes, or arrangements must appeal to truth-conduciveness' (1986: 3). Social practices are to be evaluated—*exclusively*, as far as we can tell—by how well they raise 'the *aggregate* level of [true belief] of an entire community' (1999: 93). But there are a variety of unreliable means to that end. To see why, keep in mind two facts we introduced at the outset: first, that the reliabilist evaluates beliefs, not with reference to *their* consequences, but with reference to the consequences of the types of processes that generate them; and, second, that such processes are evaluated with reference to their *direct* consequences. Elsewhere (Ahlstrom-Vij

---

[3] Specifically, we have argued (Ahlstrom-Vij 2017) that *knowing full well*, in Sosa's (2015) sense, does not add anything substantive to what Sosa in earlier work refers to as *reflective* knowledge (e.g., in his 2009), and that such knowledge in turn is reducible to *animal* knowledge, or reliable belief-formation that is creditable to the agent. Moreover, creditable epistemic success in that sense—which is also the sense we find in Greco (2010)—is reducible to belief-formation that is both reliable and safe (i.e., such that it would not easily issue in false belief, were circumstances only slightly different), which in turn is reducible to belief-formation that is reliable in a manner that's modally robust, in the sense of reliable not only in the actual world but in nearby worlds as well (Ahlstrom-Vij 2015). This means that virtue reliabilism is reducible to a modal version of the type of classic reliabilism discussed in this paper, and that it—if motivated at all—will be motivated in the terms to be discussed in what follows.

and Dunn 2014), we've put this point in terms of reliabilism being both *indirect* (in its evaluation of beliefs) and *direct* (in its evaluation of processes). While the indirect nature of reliabilism follows from our definition in RELIABILISM, its directness does not. So, a word is in order by way of motivating the latter.

Say that I'm trying to locate some particular book in my bookshelf. Some psychological process will be at work, and when evaluating the belief 'There's the book I'm looking for', the reliabilist will evaluate the reliability of that process. If the visual processes involved in locating the book are reliable, that belief will be justified. But there's a sense in which the consequences of that process extend far beyond the aforementioned visual belief. Once I've located the book, I might start reading it and form a great number of beliefs as a result. Those beliefs might lead me down a variety of different lines of inquiry, that in turn will have a multitude of doxastic consequences. But the reliabilist doesn't factor in all of *those* consequences when trying to determine whether the visual process that originally led me to the book is reliable. The only consequences relevant to the reliability of that process are the *direct* doxastic consequences of that type of process being instantiated, which, very roughly, will be evaluated in terms of the truth-ratio of the set of beliefs formed as a direct result of looking for purposes of visually locating things.

We can now see why the reliabilist is forced to accept that there are a variety of ways to raise the raise the aggregate level of true belief in a community by way of unreliable belief-formation. Consider a real-life case:

CLIMATE CHANGE: Climate change deniers believe many false things about climate change, while climate change believers believe many true things. But it's also highly likely that most deniers *and* believers have reached their respective conclusions through *politically motivated reasoning* (PMR). PMR involves 'the formation of beliefs that maintain a person's status in [an] affinity group united by shared values' (Kahan 2016: 2). Believing in man-made climate change is *not* what 'people like me' do if I'm a conservative; but it *is* what 'people like

me' do if I'm a liberal. And the distribution of belief that we see on the matter reflects this.

Is PMR *reliable*? Note two things. First, there is not a unique type of belief-forming process involved in all cases of PMR. But it's likely that there is some *family* of psychological processes involved, and that this family can be functionally defined along the lines of *believing that p because it fits with my political affiliations*. Second, the processes in this family are (in the actual world) *unreliable*. We can see this by noting that there are many different political affiliations and so different beliefs that will fit with those affiliations. It cannot be that most such beliefs are true, because many of these are in conflict with each other.[4]

Still, the processes involved in PMR can, in the right scenarios, lead to many true beliefs both directly and further downstream. This is already the case for many climate change believers—in this case, PMR has worked out well from the point of view of accuracy, on account of the reliability of those deferred to. Furthermore, work on science communication suggests that PMR can be utilized to increase the accuracy of climate change deniers, too, by offering information in a way that caters to their particular values (Kahan *et al.* 2015) or by rebranding experts in a manner that exploits deniers' trust in particular 'in group' authorities (Kahan *et al.* 2010). Indeed, we might be able to say something stronger: *using suitable packaging techniques in this manner might be the only practical way to bring deniers around to the truth about climate change*. Of course, someone might suggest that it is better to simply get people to stop relying on PMR. But notice that, even if people did stop relying on PMR, it is not clear that this would increase accuracy. These are complex and difficult issues on which lay people would tend to (and often do) go wrong, were they to coolly

---

[4] This is similar to Hume's (1975/1777) observation about miracles: Not even the religiously devout should think that the inference from 'there is testimony of a miracle' to 'there was a miracle' is a good one, because even if you think that this inference is reliable when it comes to testimony in *your* religious tradition, you should recognize that you think it fails with respect to all the other religious traditions.

examine the evidence and consult (but not necessarily defer to) the experts. That's why exploiting, not eradicating, people's reliance on PMR is likely to be more effective from the point of view of accuracy, at least in the case of climate change, even though politically motivated reasoning is not a reliable belief forming process.[5]

There are reasons to believe that this argument generalizes to other politically divisive issues (see Kahan 2010). But then a reliabilist like Goldman has a problem. In a wide variety of cases, the reliabilist is going to recommend the formation of *unjustified* beliefs in order to increase the aggregate accuracy of the community. Remember, on NORMATIVE GLOSS, unjustified beliefs are ones one should not form. So, we get the following:

> AGGREGATION: In a wide variety of cases, the reliabilist wishing to raise the aggregate level of true belief will recommend that people form beliefs in ways that they should not.

If this is correct, the very feature that enables the reliabilist to avoid intrapersonal trade-off problems—i.e., her acceptance of SIDE-CONSTRAINTS—turns out to clash with her desire to also raise the aggregate level of accuracy.[6] How might the reliabilist respond?

---

[5] Note that the family of processes involved in PMR is not changed by the fact that it is harnessed in a situation where it helps improve accuracy. It is not as if instantiating the relevant processes in a scenario where they improve people's changes of getting it right, changes the psychological processes involved into something like *politically motivated reasoning in cases where it yields true belief*. The psychological processes are what they are independently of when or for what purposes they are implemented, and hence PMR processes are still unreliable even in situations where they yield true belief.

[6] This is a good time to revisit footnote 1: We've framed our argument in terms of RELIABILISM, which holds that reliability is a necessary condition on justification. In the cases imagined, the processes involved are not reliable, so we can infer that they're unjustified. Add to this NORMATIVE GLOSS, and you get AGGREGATION.

3. The Basic Motivation for Reliabilism

Here's a natural response: AGGREGATION is not a problem for reliabil*ism*, but merely for certain reliabil*ists* who also want to evaluate social-epistemic phenomena along consequentialist lines. Reliabilism is only a view about the justification of the beliefs of individuals, and carries no implications for whether true belief should be promoted in the aggregate. So, anyone embracing RELIABILISM will not, simply on account of doing so, face AGGREGATION.

We agree. But this doesn't make the problem go away. Think about what would drive one to embrace reliabilism in the first place. We submit it's the following:

BASIC MOTIVATION: True belief is epistemically good, and should as such be promoted.[7]

Why think that BASIC MOTIVATION provides the underlying motivation for becoming a reliabilist? To start with, consider the following observation by Sanford Goldberg:

[…] reliabilist views of doxastic justification get much of their motivation from the way they honor the link between truth and justification. Belief aims at truth, and particular beliefs are justified to the extent that [they] are formed (and sustained) in such a way that they are likely to be true (Goldberg 2010: 151).

Defining justification in terms of processes that tend to issue in true belief, in the manner of Goldberg and other reliabilists, makes sense if BASIC MOTIVATION holds, since it makes clear how justification connects up with something that's good and that we want promoted. As such, it

---

[7] The reasons involved when saying here that true belief *should* be promoted are *pro tanto*. We can imagine scenarios in which non-epistemic—including moral—concerns will trump any reasons we might have for promoting true belief, for example in so far as doing so will violate people's rights (e.g., on account of privacy considerations).

also explains why we care about justification: we care about true belief, and justification helps get us (more of) something that we care about. Of course, it might be possible to motivate a broadly reliabilist account of justification in some other way—in fact, we'll look at a number of alternative motivations in Section 4—but for now we simply want to point out that BASIC MOTIVATION seems the most natural explanation.

Moreover, the naturalness of this particular motivation is increased by the fact that it also helps explain a number of *other* reliabilist projects, in addition to that of defining justification in truth-linked terms. For instance, reliabilists identify not just reliable but also *powerful* processes, or processes that generate a lot of true belief (e.g., Goldman 1992*b* and 1986); they offer advice to the individual inquirer about how to increase their chances of forming justified beliefs (Goldman 1978 and 1986, Part II), including by providing suggestions for what experts to trust (Goldman 2001) or what reasoning strategies to rely on (Bishop and Trout 2004); and they look to raise the aggregate level of true belief in society (Goldman 1999), for example by making recommendations about how to re-design epistemic environments in ways that protect people from bias (Ahlstrom-Vij 2013), or how implementing certain incentive structures might help increase the reliability of the scientific community (Kitcher 1990). Again, the most natural explanation for why the reliabilist is involved in all of these projects is that she accepts BASIC MOTIVATION.

But then we have a tension between the motivating idea behind reliabilism and the theory itself, as follows:

> BASIC TENSION: Sometimes true belief is best pursued by forming unjustified beliefs, as defined on RELIABILISM. Given NORMATIVE GLOSS, it follows from RELIABILISM that they should *not* form such beliefs; hence, we get SIDE-CONSTRAINTS. But if accept RELIABILISM on account of BASIC MOTIVATION, we think that justification is worthy of pursuit only because true belief is, which suggests that people (in at least some cases) *should* form those beliefs.

That's the deeper tension of which AGGREGATION is a symptom. Moreover, the tension puts

the reliabilist in a dilemma. She needs to either

(*a*) hold on to RELIABILISM, and with it the prohibition against unreliably formed but

truth-conducive beliefs, but then also explain why we should allow the pursuit of

justification to get in the way of acquiring true belief, while holding that justification is

valuable as a (mere) means to true belief; or

(*b*) say that it's aggregate level of true belief that matters, and then drop the prohibition

against unreliably formed but truth-conducive belief on the individual level—and in so

doing give up on RELIABILISM.

Given BASIC MOTIVATION, it's difficult to see how one could go for (*a*). One would need to

hold that true belief should be promoted, and that justification is valuable as a mere means to true

belief, but that we should in some cases pursue justification at the expense of true belief. This

would amount to a type of justification fetishism that's incompatible with also taking it to be of

(mere) instrumental value.[8] This leaves us with (*b*). Since RELIABILISM entails SIDE-

CONSTRAINTS (given NORMATIVE GLOSS), rejecting the latter is to give up on the former.

That's the sense in which any true reliabilist—as in: any reliabilist moved by the BASIC

---

[8] It won't help the reliabilist that there might be cases in which forming unjustified beliefs now will yield a great pay-off in terms of true belief later. What (*a*) requires of her is to embrace SIDE-CONSTRAINTS, and the constraints thereby embraced would under certain circumstances force her to prioritize justification over true belief even when doing so does *not* yield any accuracy pay-off further downstream. That's why talk of fetishism is called for here.

MOTIVATION—should not endorse RELIABILISM.[9]

4. Alternative Motivations for Reliabilism

We've argued that the motivation for RELIABILISM, as captured by BASIC MOTIVATION, should on reflection lead one to reject the former. But maybe there is some alternative motivation for RELIABILISM that will enable the reliabilist to block our conclusion. In this section, we'll canvas four such alternatives, and argue that they all fail.

4.1. Reliabilism and Intuitions

The reliabilist might respond that what motivates reliabilism is not the idea that what's epistemically right should be understood in terms of the epistemic good of true belief, or any desire to see more epistemically good things, but rather by the fact that that the job of the epistemologist is to generate theories that fit with our intuitions, and that reliabilism best fits our intuitions about relevant hypothetical cases.

This is not a successful response. What processes are reliable will partly be an empirical matter, which is why a reliabilist account of epistemic categories often yield counter-intuitive

---

[9] It might be objected that someone can accept BASIC MOTIVATION, and also the idea that true belief should be promoted, but deny that such promotion should take the form of raising the *aggregate* level of accuracy. Consider, for example, Coady (2010) on how the *distribution* (not merely the aggregate level) of accuracy matters in cases where we have a right to know certain things. However, note two things: First, the type of cases Coady has in mind might be best described as ones where a general, epistemic desire to raise aggregate levels is adjusted in particular cases to factor in *moral* considerations about distribution. Coady himself would deny this, but that's only because he holds that 'normative epistemology, properly understood, is a branch of ethics' (105). Second, even if we go along with Coady, there's still little comfort for the reliabilist here, as we can easily construct cases in which specific, desirable distributions of true belief are best promoted through unreliable processes—which should trouble the reliabilist no less than should AGGREGATION.

results, e.g., about reflection (Kornblith 2012), about surprisingly simple prediction rules (Bishop and Trout 2004), and about blind deference (Ahlstrom-Vij 2015). So, the claim that reliabilism best fits our intuitions about relevant hypothetical cases is false—there are alternative theories of justification that do a better job on this score. And in the absence of something like BASIC MOTIVATION, there's nothing to appeal to in suggesting that we should embrace the implications of these investigations as opposed to simply reject RELIABILISM in favor of one of these alternative theories. So, whether or not reliabilists in fact *are* motivated by BASIC MOTIVATION—and we've already argued that the claim that they are provides a unified explanation of a number of reliabilist projects—they (also) *should*.

Of course, the reliabilist can adjust her theory to fall in line with intuitions. But why would she *want* that to be the case? If reliabilism is made to be intuitive, it becomes unmotivated. Again, whether or not beliefs are justified will be hostage to empirical circumstances. We cannot assume that these circumstances will always (or even generally) be reflected in our intuitions about cases. And notice that no *general* skepticism about the reliance on intuitions in philosophy is required to make that claim. Independently of whether we are attracted to reliabilism or not, we have reason to believe that our intuitions might not track the empirical facts that help determine the reliability of any given process, including (or perhaps especially) in hypothetical circumstances, if only because these facts will often be surprising and far from obvious. Consequently, if you're a reliabilist, there's no good reason for seeking a particularly close fit between theory and intuitions.

4.2. Reliabilism and Naturalistic Intuition Satisfying

So, the motivation for reliabilism cannot simply be that it best captures our intuitions about justification. But there is a more nuanced motivation for reliabilism, which might seem to fare better. This motivation comes from Goldman (1979). According to this motivation, we should adopt a theory of justification that (*i*) is naturalistically acceptable and (*ii*) best captures our intuitions about justification. Goldman's basic idea is that a theory of justification is naturalistically

acceptable if it makes no use of evaluative or deontic terms or concepts in its statement (90). Reliabilism seems to fit the bill here, since justification is specified in terms such as 'psychological process' and 'ratio of true beliefs to false beliefs', none of which are evaluative epistemic terms.

Goldman doesn't say much by way of defending the naturalistic constraint. But Kornblith (2018) offers a plausible motivation with reference to what Fodor says about intentionality, when he notes that '[i]f aboutness is real, it must be really something else' (1987: 97). According to Kornblith, something similar is going on in the case of Goldman on justification:

> Being justified is not one of the fundamental properties of things, in the way that, perhaps, spin, charge, and charm are. Nor are any of the other epistemic properties we might explain justification in terms of. What this means, however, is that any account of justification which locates it squarely within the natural world—that is, the only world there is—must explain justification in non-epistemic terms. To paraphrase Fodor: If epistemic properties are real, they must really be something else (Kornblith 2018: 27).

The idea, then, is that we have certain intuitions about justification: my belief that I have hands is justified, beliefs in the predictions of astrology are unjustified, and so forth. Reliabilism about justification vindicates these intuitions and yet is naturalistically acceptable, in that it picks out the natural properties that distinguish the justified from what's not. Other theories of justification may do better in terms of capturing (a wider set of) intuitions, but, the thought goes, they are not naturalistically acceptable. Reliabilism is the unique theory that does best with respect to (*ii*) while satisfying (*i*).

One way to see the problem with this motivation for reliabilism, however, is to point out how odd the project starts to look once we consider cases like those that occupy Goldman (1979) towards the end of section II of his paper. After presenting his reliabilist theory of justification, he considers what to say about processes of belief formation, like wishful thinking, that though

unreliable in our world may be reliable in other possible worlds. Goldman is indecisive in the face of such counterexamples. In the end, he suggests the following:

> What we really want is an *explanation* of why we count, or would count, certain beliefs as justified and others as unjustified. Such an explanation must refer to our *beliefs* about reliability, not to the actual *facts*. The reason we *count* beliefs as justified is that they are formed by what we *believe* to be reliable belief-forming processes. […] What matters, then, is what we *believe* about wishful thinking, not what is *true* (in the long run) about wishful thinking (Goldman 1979: 101).

This is a puzzling passage. If all we are we are looking to do is determine what we *believe* about this or that particular way of forming beliefs, there is no need for a naturalistic constraint. For in that case, we are presumably simply mapping out our intuitions about cases. The naturalistic constraint only makes sense if we read Goldman in the way that Kornblith does, as being in the business of telling us what justification *is*, as in: what real features of the world distinguish what's justified from what's not. Of course, you might think that, by mapping out our intuitions, we *in effect* find out what justification (really) is. But if you to think that, and you are a reliabilist, then you would also have to think something along the lines of the following: once we list all of the intuitive ways of forming beliefs justifiably, and all of the intuitive ways of forming beliefs unjustifiably, we'll find that everything in the first column is intuitively reliable, and everything in the second column is intuitively unreliable. But why think *that*? Every single intuitive counterexample to reliabilism—including those involving evil demons, reliable clairvoyants, and so forth—should lead us to expect otherwise.

This is not to suggest that the reliabilist should cease to be a reliabilist on account of these counterexamples; it's to say that reliabilists should not be heavily invested in the business of trying to accommodate intuitions about justification in the first place. Better then to drop the

requirement that our account of justification should provide the best fit with our intuitions about cases, while holding on to the idea that we should try to understand what justification really is. That's where the naturalistic constraint fits in: as Kornblith puts it, if epistemic properties are real, they must really be something else. But now we have a problem: if all we want is a naturalistically acceptable story, there will be plenty to choose from. How are we to narrow them down to arrive at a theory of justification? To start with, by invoking BASIC MOTIVATION. That would narrow down the field to those naturalistically acceptable stories that (as Goldberg put it) honor the link between justification and true belief. The most established story that fits the bill is RELIABILISM. But if what motivates the latter is BASIC MOTIVATION, we are of course back with the problem we started out with—which demonstrates that the attempt at an alternative motivation that can be found in Goldman (1979) ultimately offers no such thing.

4.3. Consequentialist Intuition Satisfying

In Goldman (1986), we get a different motivation for reliabilism, and also end up with a view that is subtly different from the one proposed in Goldman (1979). In his (1986), Goldman explicitly describes his approach as one of reflective equilibrium (60) according to which intuitions about justification have an important role to play. And he continues to want a theory of justification on which justification is defined using non-evaluative terms, though this plays a less prominent role. But there is also a new ingredient, which is an explicit commitment to some form of epistemic consequentialism. As Goldman puts it, he wants a theory of justification that is 'truth-linked' (69). More specifically, '[t]rue belief is the value that J-rules [rules dictating which beliefs are justified] should promote—really promote—if they are to qualify as right' (103).

These comments about epistemic consequentialism seem to go *against* a reliabilist account of justification, however, and instead in favor of an account that evaluates each belief, not with reference to its etiology, but in terms of its epistemic consequences. Why, then, does Goldman (1986) reject such a view?

Goldman doesn't say[10], but we suspect that the reason has to do with the fact that the theory we are left with does well with respect to its consequentialist credentials, but fares poorly with respect to our intuitions about justification. In a telling quote, Goldman considers proposals to 'regiment' the concept of justification in various ways that makes it more theoretically pleasing (it doesn't matter for our purposes what these regimented proposals look like). He writes: 'Either of these [regimented] approaches might seem preferable from a systematic or theoretical point of view. Nonetheless, they do not seem to be what is implied by the ordinary conception [of justification] as it stands; and that is all I am currently trying to capture' (109).

So, in Goldman (1986) we seem to be getting the following kind of rationale for reliabilism: it is the theory that (*i*) is in some sense consequentialist in that it promotes true belief, and yet (*ii*) also satisfies our intuitions about justification. Three problems now present themselves.

First, though he relies heavily on intuition-satisfying, he also says things that cut against it. For instance, in dismissing a theory of justification according to which a belief is justified just in case it is in conformity with the belief-forming rules accepted by one's society, Goldman writes: 'Any such proposal invites an obvious objection. Why should we assume that what is accepted as justification-conferring by the members of a particular community really is justification-conferring? Can't such a community be wrong?' (68). We agree with Goldman that the answer here is 'yes.' But then we have a problem for (*ii*) above, since what holds of a community's beliefs about justification surely also holds of our *intuitions* about it, since these too would have to be defined with reference to some community or other. Goldman could of course drop (*ii*) and rely entirely on (*i*). Given what we have argued, the most natural way to proceed from this point would be by appealing to BASIC MOTIVATION—but then we're back with the problem that we started out with.

---

[10] Goldman's explicit reason is not explanatory: 'I ignore entirely the suggestion that the justificational status of each belief is a function of that very belief's consequences' (1986: 97).

Second, in explaining why he doesn't go for a coherence-based account of justification, Goldman writes: 'The fundamental standard concerns the formation of true belief. Coherence enters the picture only because coherence considerations are generally helpful in promoting true belief' (100). Why not say the same thing about reliably formed belief? Reliably formed beliefs usually promote true belief, but in some cases—such as in CLIMATE CHANGE—unreliably formed beliefs do. So, our criterion of what makes a belief epistemically right is not that it coheres with other beliefs, nor is it that it is reliably produced. That's to say that, if we go down this route, we have not found a *motivation* for reliabilism that's separate from what might or might not be offered by BASIC MOTIVATION; if anything, we have found a reason to *reject* reliabilism on grounds consistent with BASIC MOTIVATION.

Third, Goldman's preferred criterion of rightness actually seems to permit unreliably formed beliefs as justified. Here it is:

(ARI) A J-rule system R is right if and only if R permits certain (basic) psychological processes, and the instantiation of these processes would result in a truth ratio of beliefs that meets some specified high threshold (greater than .50).

Though this has not been widely noted[11], it is certainly possible that some set of psychological processes yields a truth ratio of beliefs above the threshold and yet one of the psychological processes is itself unreliable. So, Goldman's own arguments against other views, and indeed his official statement of his theory, seems to give significant weight to the truth-linkedness motivation for reliabilism. And that is of course exactly as it should be, if BASIC MOTIVATION is what is driving the project. But as we have seen, this is also the very motivation that ultimately undermines reliabilism.

---

[11] See Dunn (2012) for a discussion.

4.4. Knowledge as a Natural Kind

So far we have focused on Goldman, and found that we run into problems when attempting to defend a view that honors some form of consequentialism and also best captures our intuitions. So, let us consider a view that depends on neither of these claims: Kornblith's (2002) proposal to investigate knowledge as a natural kind. On Kornblith's view, the proper subject matter of epistemology does not consist of our epistemic concepts, but rather of *knowledge itself*. According to Kornblith, we can study knowledge itself by looking at the attributions of knowledge that biologists and cognitive ethologists are required to make to explain the survival and evolution of animals with robust cognitive systems. Here is Kornblith, summing up where he thinks this approach takes us:

> The knowledge that members of a species embody is the locus of a homeostatic cluster of properties: true beliefs that are reliably produced, that are instrumental in the production of behavior successful in meeting biological needs and thereby implicated in the Darwinian explanation of the selective retention of traits. The various information-processing capacities and information-gathering abilities that animals possess are attuned to the animals' environment by natural selection, and it is thus that category of beliefs that manifest such attunement—cases of knowledge—are rightly seen as a natural category, a natural kind (Kornblith 2002: 62-3).

On this view, we get a naturalistic, non-intuition-based motivation for the claim that knowledge is reliably produced true belief. How could one leverage this into a non-intuition-based motivation for reliabilism about justification? The natural way to go is to claim that justification just is whatever we add to true belief to get knowledge. As such, justification would form a proper part of the homeostatic property cluster that is knowledge.

In evaluating this approach, let us grant Kornblith's (controversial) claim that knowledge is a natural kind, which just is reliably produced true belief. Even granting this, there are problems in getting a motivation for reliabilism about justification out of his view. To begin with, there is a fairly wide consensus that adding justification to true belief does *not* get you knowledge. Kornblith might respond that this consensus is based on a type of intuition-based conceptual analysis that he rejects. The problem for Kornblith (were he to attempt a motivation for reliabilism along the lines we are considering) is that the same would have to go for the claim that justification just *is* what you add to true belief to get knowledge. That claim, too, is a kind of intuition-backed claim of the sort we're supposed to be disavowing here.

Of course, Kornblith might deny that this claim is—or at least: *needs* to be—motivated with reference to intuitions. Remember, the idea we're considering is that justification is a proper part of the homeostatic property cluster that is knowledge. So, perhaps a reliabilist analysis of justification simply falls out of Kornblith's account of knowledge. Or does it? What's providing the explanation of the selective retention of traits, and thereby warranting the talk of a natural kind, is the (whole) property cluster that—following cognitive ethologists—we can call 'knowledge,' not any proper part of it. If any proper part of it *did* provide the relevant explanation, then *that* would be the natural kind in question. So, as far as Kornblith's account goes, we have a motivation for slicing the world into what's known and not known, but not for slicing the world into what's justified and not justified. That's not to say that Kornblith's account *rules out* there being such a motivation; it's just to say that it doesn't provide one.

But maybe there's a motivation to be had by looking at what Kornblith has to say about epistemic normativity more generally. According to Kornblith, 'someone who cares about acting in a way that furthers the things he cares about, and that includes all of us, has pragmatic reasons to favor a cognitive system that is effective in generating truths, whether he otherwise cares about the truth' (2002: 156). For that reason, Kornblith suggests, '[w]e should […] adopt a method of cognitive evaluation that endorses truth-conducive processes' (156). If Kornblith is right, we have

a pragmatic reason to slice the world into the truth-conducive and the non-truth-conducive. But there are two problems here:

The first problem is that Kornblith is likely wrong. As argued by Hazlett (2013), there is a 'clearly identifiable pattern of cases in which false belief is better for a person than true belief' (63), on account of how self-serving illusions promote non-depression and self-esteem, in turn conducive to subjective well-being and the ability to cope. These are things that we want—indeed, possibly the things we want *the most*—which suggests that the relationship between true belief and desire satisfaction is not as straightforward as Kornblith suggests.[12]

The second problem is that, even if Kornblith is right, it doesn't help the reliabilist. It doesn't follow from the fact—if it *is* a fact—that we all have pragmatic reason to value truth that we should 'adopt a method of cognitive evaluation that endorses truth-conducive processes' (156). We've already seen why: under many circumstances—including those in CLIMATE CHANGE—the best way to promote true belief is *not* to endorse truth-conducive processes. Of course, Kornblith might object that we are here assuming that the processes involved are *belief-forming* processes. Perhaps the relevant processes here are instead the *social* processes that involve applying some particular accuracy-motivated intervention on a social group, such as one exploiting politically motivated reasoning. And if that raises the aggregate, and to that extent promotes what we all have reason to care about on Kornblith's view, then maybe *that's* the type of process we should endorse.

---

[12] It might be suggested that Hazlett's observation thereby also calls into question BASIC MOTIVATION. In particular, if Hazlett is right, in what sense *should* true belief be promoted? One answer consistent with what's been said in this paper—although we do not for present purposes need to commit to it—can actually be found in Hazlett's book. Hazlett argues that evaluative practices take place within a 'critical domain' organized around the value of true belief. To the extent that we care about believing truly under some particular circumstances, epistemic evaluation will be relevant to us, and the 'should' in BASIC MOTIVATION will have force. At other times, non-epistemic considerations will carry greater weight—all of which accords well with the fact, highlighted in footnote 7, that the 'should' in BASIC MOTIVATION is *pro tanto*.

But if Kornblith says that, then that merely highlights the fact that his account doesn't motivate *reliabilism* after all—remember, the reliabilist would have to say that the beliefs involved in these cases are ones people should not form—whether independently of BASIC MOTIVATION or not. If anything, this offers a motivation for *giving up* on reliabilism that's independent of BASIC MOTIVATION.

4.5 Cognitive Contact with Reality

We will consider one final motivation for reliabilism that attempts to evade the dilemma we've presented here for the reliabilist.[13] According to this view, what is epistemically valuable is a kind of cognitive contact with reality. Something like this, for instance, is what distinguishes lucky guesses from knowledge. So what is it to have cognitive contact with reality? Perhaps, the thought goes, we can understand this in terms of reliability. When I *know* that the bottle is on the table (and don't merely guess correctly that it is), it is because there is some kind of reliable connection between the fact that it is on the table and my cognitive state.

Such a motivation for reliabilism avoids the dilemma we've suggested as follows. You shouldn't hold unreliably formed but truth-conducive beliefs because such beliefs wouldn't be justified. And justification can get in the way of the pursuit of truth because justification (that is: cognitive contact with reality) is valuable in itself. It is not, then, just the pursuit of true belief that matters, because cognitive contact with reality is important in its own right.

Though this view may escape the dilemma as initially stated, we think it nevertheless runs into a closely related one. To see this, note that there are two ways to understand the claim that reliability gives one cognitive contact with reality. First, one might say that, whenever one has a reliably formed belief, one has cognitive contact with reality. But that view isn't plausible, we think, because there can be reliably formed beliefs that are false. When one has a reliably formed but

---

[13] Thanks to an anonymous reviewer from this journal for prompting us to think about this kind of view.

false belief, one surely doesn't have any contact with reality.

The other option, then, is that one has cognitive contact with reality when one has a reliably formed *true belief*. But if that is what is valuable, then surely you should take the following trade-off: form a belief via an unreliable process now that leads to many more reliably produced true beliefs later. The dilemma, then, isn't really escaped. Such a reliabilist will hold on to reliabilism's prohibition against unreliably formed beliefs. But she will have to do so even when these unreliably formed beliefs are conducive to giving the believer cognitive contact with reality. In so doing, she must explain why we should allow the pursuit of justification to get in the way of the pursuit of cognitive contact with reality, while holding that justification is valuable just because it gives one cognitive contact with reality.

5. Advice for Former Reliabilists

Our aim thus far has been a negative one: we have offered an argument to convince true reliabilists—those who embrace RELIABILISM because they buy into BASIC MOTIVATION—*not* to endorse RELIABILISM. But what are former reliabilists to do if we're right?

5.1. *Epistemic Utilitarianism*

We can't give anything like a complete answer to this question here. But the broad outline of the view that former reliabilists should pursue is already clear. It's the type of view relied on by Goldman in his social epistemology, when suggesting that social practices are to be evaluated by how well they raise 'the aggregate level of [true belief] of an entire community' (1999: 93). Such evaluation is easy to motivate against the background of BASIC MOTIVATION. But what's missing if we surrender RELIABILISM is a framework for evaluating *individual* agents. The idea of evaluating social practices with reference to aggregate true belief doesn't give us *that*, or at least not directly. By way of concluding this paper, we will therefore offer some advice to former reliabilists about how to extract such a framework from Goldman's idea about evaluating aggregate

levels of accuracy.

Let's start by considering some desiderata. Since we are assuming that the former reliabilist is motivated by BASIC MOTIVATION, we need a framework that's motivated by the idea that true belief is epistemically good and should as such be promoted. Moreover, what we learn from AGGREGATION is that, in at least some cases, true belief is best promoted by forming unjustified beliefs. So, we need a view that will recommend individuals forming beliefs by unreliable methods in cases where so doing increases the aggregate level of true belief. We believe the view that fits the bill here is a form of *epistemic utilitarianism*, on which epistemic agents should form beliefs in such a way as to promote true belief across an entire epistemic community—even if that means forming beliefs by way of unreliable belief-forming processes. Such a view is easy to motivate with reference to BASIC MOTIVATION, and also takes the lesson of BASIC TENSION to heart.

5.2. Against Epistemic Egoism

However, the resulting view might be taken to be obviously incorrect. Consider a recent case offered by Julia Driver (2018):

> Conrad is the sort of person who consistently offers very bad arguments for his beliefs. Almost all of his beliefs are false, and yet he manages, amazingly, to survive. His arguments are so bad, in fact, that whenever someone talks to Conrad they end up, for very good reason, believing the opposite of Conrad's conclusions. Since Conrad's conclusions are invariably false, his interlocutors always come away with true beliefs (Driver 2018: 119).

Driver goes on to suggest that '[o]ne is better off epistemically, if one deviates from Conrad's set of beliefs. It would be very odd, then to view Conrad as a good epistemic agent in virtue of the fact that wherever he goes, true beliefs are produced' (119). The idea is that Conrad isn't doing

well epistemically even though he brings about an abundance of true belief in, since none of those beliefs are good for Conrad.

We think that this is correct. And it might be taken to suggest the following argument: if Conrad isn't doing *well* epistemically, he's not doing what he (epistemically) *should* be doing; hence, acknowledging that Conrad isn't doing well is incompatible with epistemic utilitarianism. But that's not right. What Driver's case makes plausible is that epistemic goods are what Smith (2003) refers to as *relative* goods. Goodness is relative in this way just in case its existence presupposes a relation between something (in our case, true belief) and a particular person (say, Conrad). On this picture, there is not just this good thing (a true belief) that Conrad happens to have. Rather, the epistemic goodness is inherent in the relation between Conrad and his true belief: his true belief is good-for-Conrad.[14] If that's so, we can account for why it seems wrong to say about Conrad that he's doing well, just because he's the cause of many true beliefs; since (let's assume) none of those true beliefs belong to him, they are not good-for-him and he's not doing epistemically well.

The reason that this doesn't generate a problem for the epistemic utilitarian is that we don't necessarily settle the question of what you *should* do by finding out what's *good* for you. It might be that you sometimes should do things that aren't good for you. Of course, in ethics, this is exactly the claim denied by the ethical egoist, so the view the epistemic utilitarian needs to show implausible here would be an analogous form of *epistemic* egoism, on which you should only promote what's epistemically good for you. We think there are reasons to reject that position, and to deny in particular that the only thing that an agent is epistemically obligated to promote are those things that are epistemically good for that agent.

---

[14] Smith introduces relative goods as a way to reject Moore's (1988/1902: sec. 59) argument to the effect that ethical egoism is self-defeating. It is slightly surprising, perhaps, that we—as will be clear in a moment—want to make use of the same concept to defend epistemic utilitarianism against epistemic egoism. We explain this more below.

To begin with, to claim that an agent is only epistemically obligated to promote what is epistemically good for her would be to claim that there is something *special* about one's own good, as far as what deserves to be promoted. But this is *ad hoc*. To assign such special status to one's own good would be to deny Sidgwick's (1981/1874) second axiom about goodness: 'as a rational being I am bound to aim at good generally—so far as it is attainable by my efforts—not at a particular part of it' (382). To aim only at what is good-for-me is to aim at only a particular part of the good—which raises a question: why only aim at *that* part?

Smith (2003) has an answer when it comes to his preferred account of *moral* goodness. It appeals to his particular dispositional theory of moral value. For Smith, when someone judges *p*'s being the case to be good-for-her what she believes is that she would desire that *p* if she had a maximally informed and coherent and unified set of desires. So, if I judge that *p* is good-for-me, it is built in that I desire that *p*. But if I judge that *p* is good-for-you, I believe that if you had a maximally informed and coherent and unified set of desires then you would desire that *p*. My judging that *p* is good-for-me, then, makes a claim on me (via my desires) that my judging that *p* is good-for-you does not, since the latter is about what I think about your desires, and not about my own. For Smith, if something does not make a claim on me it cannot generate obligations for me. So, he has a story about why I am obligated to promote what is good-for-me, but not what is good-for-you.

But absent any reason to posit such a dispositional theory about *epistemic* value, this kind of response doesn't work and the *ad hoc* worry raised above still stands. And so long as this *ad hoc* worry stands, the operative norm for the reliabilist is the one given by BASIC MOTIVATION: that true belief is good, and as such should be promoted. That some goods are good-for-you while others are good-for-others doesn't make a difference as such to what you should promote, and there will be at least some cases where your reasons to promote some goods over others work out in such a way that you are obligated to promote what's good-for-others even if that's not good-for-you. But if that's so, then epistemic egoism is false.

5.3. Interpersonal Trade-offs

Of course, the fact that the epistemic egoist's approach to epistemic obligation is mistaken does not entail that epistemic utilitarianism is true, or even that it's in the final analysis plausible. In order to defend a fully utilitarian view in epistemology, we would have to argue that, in any situation, epistemic agents epistemically ought to promote the aggregate epistemic good. By way of illustration, return to Conrad. Is he doing well epistemically? No, and the reason is that epistemic goods are relative goods. Is he *obligated* to be the highly effective anti-expert that he is? On epistemic utilitarianism, he might be. Just consider all of the epistemic good he generates. Surely, these give him *some* epistemic reason to be the anti-expert that he is.[15]

Of course, having *reason* to do something doesn't mean you're *obligated* to do it. Depending on how we imagine the broader context in which Conrad operates, judgments about his being obligated to be the type of anti-expert he is become more or less plausible. Say, for example, that all of his bad arguments can be traced back to one fundamental misunderstanding that is fairly easy to correct. If so, there's a nearby possible world in which Conrad offers *good* arguments, and as such makes a valuable epistemic contribution to his community, without having to be as badly off epistemically as in the world Driver imagines. If that's so, maybe what he should be doing is to correct that misunderstanding. However, if we assume that any (sufficiently) nearby possible world in which he's not an anti-expert is epistemically inferior to the one in which he is by a wide enough margin—the aggregate level of true belief in these alternative worlds are substantially deflated, say—then an epistemic utilitarian will say that Conrad should do what he's doing, even if he is thereby personally doing epistemically very poorly. (This is compatible with there being

---

[15] And, as argued above, that point holds even if all epistemic goods are relative goods; it simply doesn't follow from all epistemic goods being relative goods that each agent only has reason to promote what's good-for-her.

ample—and overriding—*moral* considerations against that arrangement, but that's a matter for ethics, not epistemology, to arbitrate.)

What this gets to is the possibility of *interpersonal* epistemic trade: in at least some cases, promoting the aggregate level of true belief calls for trading off the epistemic welfare of some for the benefit of raising that aggregate level. That possibility makes sense, given the trajectory of the reliabilist's chain of commitments. Again, reliabilists avoid intrapersonal trade-offs because RELIABILISM entails SIDE-CONSTRAINTS, but the latter sits uncomfortably with her commitment to BASIC MOTIVATION and should therefore be rejected. Instead, the way for former reliabilists to honor BASIC MOTIVATION is to hold, with the epistemic utilitarian, that agents should promote the aggregate level of true belief—including in cases where that's not epistemically good for them (or: good-for-them). Hence, the possibility of interpersonal trade-offs sanctioned by the epistemic utilitarian.

Fully pursuing the implications of this utilitarian framework goes beyond the scope of this paper, as it would involve delicate questions about how to think about the aggregate epistemic good and the specific merits of different trade-offs. What seems clear, however, is that true reliabilists should surrender reliabilism for epistemic utilitarianism. More specifically, any plausible form of epistemic utilitarianism needs to hold that epistemic goods are relative goods—only then will they be able to make plausible judgments about how well individual agents are doing—and individual agents should believe in ways that will promote the aggregate level of true belief in her community, understood along the lines of Goldman's framework for evaluating social institutions from an epistemic point of view. Making these points doesn't show that epistemic utilitarianism is *true*. But it provides the direction that former reliabilists should go.

5. Conclusion

Our main aim in this paper has been to offer an argument to convince true reliabilists—those who embrace RELIABILISM because they buy into BASIC MOTIVATION—*not* to endorse

RELIABILISM. Additionally, we have offered some advice for former reliabilists, in suggesting that they should pursue a form of epistemic utilitarianism, on which we sometimes should promote the good of others, including our epistemic community as a whole. We have not, however, offered a full-blown defense of epistemic utilitarianism, a crucial component of which would involve a mapping out of the type of interpersonal trade-offs that would be called for on that view. Still, it's our hope that what we have done here helps motivate a sustained and systematic investigation into the merits of epistemic utilitarianism in potentially dislodging reliabilism from its position as one of the most influential views in epistemology over the past couple of decades.

References

Ahlstrom-Vij, K. (2013). *Epistemic Paternalism: A Defence*, Basingstoke: Palgrave Macmillan.

Ahlstrom-Vij, K. (2015). 'The Social Virtue of Blind Deference,' Philosophy and Phenomenological Research 91(3): 545-582.

Ahlstrom-Vij, K. (2017). '*Judgment and Agency*, by Ernest Sosa,' *Australasian Journal of Philosophy* 95(1): 196-199.

Ahlstrom-Vij, K., and J. Dunn (2014). 'A Defence of Epistemic Consequentialism,' *Philosophical Quarterly* 64(257): 541-551.

Berker, S. (2013*a*). 'Epistemic Teleology and the Separateness of Propositions,' *Philosophical Review* 122(3).

Berker, S. (2013*b*). 'The Rejection of Epistemic Consequentialism,' *Philosophical Issues* 23.

Bishop, M. and J. D. Trout (2004). *Epistemology and the Psychology of Human Judgment*. Oxford: Oxford University Press.

Coady, D. (2010). 'Two Concepts of Epistemic Injustice,' *Episteme* 7(2), pp. 101-113.

Driver, J. (2018). 'The "Consequentialism" in "Epistemic Consequentialism",' in *Epistemic Consequentialism*, eds, K. Ahlstrom-Vij and J. Dunn, Oxford: Oxford University Press.

Dunn, J. (2012). 'Reliabilism: Holistic or Simple?' *Episteme* 9(3): 225-233.

Dunn, J. and K. Ahlstrom-Vij (2017). 'Is Reliabilism a Form of Consequentialism?' American Philosophical Quarterly 54(2): 183-194.

Firth, R. (1981). 'Epistemic Merit, Intrinsic and Instrumental,' *Proceedings and Addresses of the American Philosophical Association* 55(1)

Fodor, J. (1987). *Psychosemantics: The Problem of Meaning in the Philosophy of Mind*. Cambridge, MA: MIT Press.

Goldberg, S. (2010). *Relying on Others*. Oxford: Oxford University Press.

Goldman, A. (1978). 'Epistemics: The Regulative Theory of Cognition,' *Journal of Philosophy* 75(10).

Goldman, A. (1979). 'What Is Justified Belief?' in *Justification and Knowledge*, ed., G. Pappas, Dordrecht, Holland: Springer.

Goldman, A. (1986). *Epistemology and Cognition.* Cambridge, MA: Harvard University Press.

Goldman, A. (1992*a*). 'Foundations of Social Epistemic,' in his *Liaisons: Philosophy Meets the Cognitive Sciences.* Cambridge, MA: MIT Press.

Goldman, A. (1992*b*). 'Epistemic Folkways and Scientific Epistemology,' in his *Liaisons: Philosophy Meets the Cognitive Sciences.* Cambridge, MA: MIT Press.

Goldman, A. (1992*c*). 'Epistemic Folkways and Scientific Epistemology,' in his *Liaisons: Philosophy Meets the Cognitive Sciences.* Cambridge, MA: MIT Press.

Goldman, A. (1999). *Knowledge in a Social World.* Oxford: Oxford University Press.

Goldman, A. (2001) 'Experts: Which Ones Should You Trust?' *Philosophy and Phenomenological Research* 63(1): 85-110.

Greaves, H. (2013). 'Epistemic Decision Theory,' *Mind* 122(488).

Greco, J. (2010), *Achieving Knowledge: A Virtue-Theoretic Account of Epistemic Normativity.* Cambridge: Cambridge University Press.

Hazlett, A. (2013). *A Luxury of the Understanding: On the Value of True Belief.* Oxford: Oxford University Press.

Hume, D. (1975). *An Enquiry Concerning Human Understanding*, 3rd edition, ed., Selby-Bigge and Nidditch, Clarendon Press: Oxford; originally published in 1777.

Jenkins, C. (2007). 'Entitlement and Rationality,' *Synthese* 157(1).

Kahan, D. (2010). 'Fixing the Communications Failure,' *Nature* 463.

Kahan, D. (2016). 'The Politically Motivated Reasoning Paradigm, Part 1: What Politically Motivated Reasoning Is and How to Measure It,' *Emerging Trends in the Social and Behavioral Sciences*, 1-16.

Kahan, D. H. Jenkins-Smith, T. Tarantola, C. L. Silva, and D. Braman (2015). 'Geoengineering and Climate Change Polarization: Testing a Two-Channel Model of Science Communication,' *The Annals of the American Academy of Political and Social Science* 658(1).

Kahan, D., D. Braman, G. L. Cohen, J. Gastil, and P. Slovic (2010). 'Who Fears the HPV Vaccine, Who Doesn't, and Why? An Experimental Study of the Mechanisms of Cultural Cognition,' *Law and Human Behavior* 34.

Kitcher, P. (1990). 'The Division of Cognitive Labor,' *The Journal of Philosophy* 87(1), pp. 5-22.

Kornblith, H. (2002). *Knowledge and Its Place in Nature.* Oxford: Oxford University Press.

Kornblith, H. (2012). *On Reflection.* Oxford: Oxford University Press.

Kornblith, H. (2018). 'The Naturalistic Origins of Epistemic Consequentialism,' in K. Ahlstrom-Vij and J. Dunn, eds, *Epistemic Consequentialism*. Oxford: Oxford University Press.

List, C. and P. Pettit (2004). 'An Epistemic Free-riding Problem?' in P. Catton and G. Macdonals (eds.), *Karl Popper: Critical Appraisals*, Routledge, pp. 128-158.

Littlejohn, C. (2012). *Justification and the Truth Connection*. Cambridge: Cambridge University Press.

Moore, G. E. (1988). *Principia Ethica*, Amherst, NY: Prometheus Press; originally published in 1902.

Nozick, R. (1981). *Philosophical Explanations*. Cambridge, MA: Belknap Press.

Rawls, J. (1999). *A Theory of Justice, Revised Edition*. Cambridge, MA: Harvard University Press.

Sidgwick, H. (1981). *The Methods of Ethics*, 7th edition. Hackett Publishing Company; originally published in 1874.

Smith, M. (2003). 'Neutral and Relative Value after Moore,' *Ethics*, 113: 576–98.

Sosa, E. (2009). *Reflective Knowledge: Apt Belief and Reflective Knowledge, Vol. II*. Oxford: Oxford University Press.

Sosa, E. (2015). *Judgment and Agency*. Oxford: Oxford University Press.

Thomson, J. J. (1976). 'Killing, Letting Die, and the Trolley Problem,' *Monist* (59).