Episteme (2026), **23**, 388–410
doi:10.1017/epi.2024.54

episteme

ARTICLE

On the value of changing your mind

Matthew Shields

Wake Forest University, Winston-Salem, NC, USA
Email: shieldm@wfu.edu

(Received 18 March 2024; revised 16 August 2024; accepted 14 September 2024; first published online 27 January 2025)

Abstract

We are all capable of arriving at views that are driven by corrupting non-epistemic interests. But we are nonetheless very skilled at performing a commitment to epistemic goods in such cases. I call this the "Problem of Mere Epistemic Performance," and it generates a need to determine when these commitments are illusory and when they are in fact genuine. I argue that changing one's mind, when done in response to the evidence and at a likely cost to oneself, is the best indication that an agent is committed to epistemic goods and that they are genuinely in the game of giving and asking for reasons. This is because changing one's mind in this way goes as far as we can in eliminating the possibility that the agent has an ulterior motivation for their epistemic practices. Moreover, this account shows that the consensus view of the ideal epistemic agent is mistaken. The ideal agent *must* have false beliefs or deficient epistemic practices because only then do they have the opportunity to change their mind and establish a commitment to epistemic goods – a commitment that even an agent with only true beliefs and maximal justification or understanding may lack.

Keywords: Epistemic goods; virtue epistemology; commitment; ideal and non-ideal epistemology

1. Introduction

In the eulogies for Hilary Putnam, we find the common refrain that one of Putnam's most admirable qualities as a thinker was his willingness to change his mind – a quality, it is emphasized, few philosophers possess. Here, for example, is Martha Nussbaum on Putnam: "Most philosophers talk a lot of talk about following the argument, but eventually lapse into dogmatism, defending a well-known position at all costs, no matter what new argument comes along. The glory of Putnam's way of philosophizing was its total vulnerability" (Nussbaum 2016). Nussbaum's praise goes further still: "What Hilary Putnam's life offers our troubled nation is, I think, a noble paradigm of a perpetual willingness to subject oneself to reason's critique" (2016).

Putnam himself was aware of this reputation, and he cites his own philosophical inspiration for this approach: "Carnap – a great philosopher who had an aura of integrity and seriousness which was almost overwhelming – would stress that he had changed his mind on philosophical issues, and changed it more than once. 'I used to think . . . I *now* think' was a sentence construction that was ever on Carnap's lips . . . Carnap is still the outstanding example of a human being who puts the search for truth higher than personal vanity" (1988: xii).

© The Author(s), 2025. Published by Cambridge University Press. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted re-use, distribution and reproduction, provided the original article is properly cited.

https://doi.org/10.1017/epi.2024.54 Published online by Cambridge University Press

Examples of this effusive praise for the practice of changing one's mind are certainly not limited to philosophy. Diane Ravitch, a preeminent historian of education and Assistant Secretary of Education under George H.W. Bush, spent a great deal of her academic career and time with the Bush administration advocating for increasing standardized testing and the privatization of public education. Observing these policies play out, however, Ravitch radically changed her views, becoming a vehement critic of both the George W. Bush and the Obama administrations' continued advocacy of this approach and instead a staunch proponent of public schools and teachers' unions. In *The New York Times* review of Ravitch's book documenting this shift – after noting that "Ravitch's offer to guide us through this mess comes with a catch: she has changed her mind" – we find the following claims: "[W]e are lucky to have Ravitch as" our teacher, and while "I wish we could all share Ravitch's open-mindedness in seeing what the data really tells us. Somehow, I doubt that's what will carry the day" (Wolfe 2010).[1] Peter Beinart, a prominent journalist and academic, has changed his views on Israel–Palestine substantially over time, now advocating for a one-state solution after being a vocal advocate of a two-state approach. While receiving intense criticism from various quarters, he has also received a great deal of praise – for example, that he is willing "to look honestly and bravely at reality" and that "Beinart is himself a source of pride" (Levy 2020).

Or consider this assessment of the case of evolutionary biologist Daniel Bolnick, who realized a striking result he had published was the result of faulty coding and contacted the publisher to retract the paper. A journalist writes of this case: "I think society ought to give Bolnick some sort of a prize . . . We need more Heroes of Retraction" (Thompson 2018). Carlo Rovelli explains his admiration of Einstein in a similar vein. Few scientists "have changed their minds as frequently as he did" (2020: 77). Einstein's changes of mind were the product of "genuine scientific errors: mistaken ideas, wrong predictions, error-strewn equations, scientific assertions that he himself came to regret and that were proved false," including claims concerning the expansion of the universe, the existence of gravitational waves, and the nature of spacetime (77). "Do all these mistakes and changes of opinion take something away from our admiration of Albert Einstein? Not at all. If anything, the opposite is the case" (80). Einstein's willingness "to make mistakes, to change one's ideas, not once but repeatedly" are traits to be admired (81).[2]

When we praise a change of mind in these contexts, we of course do not just mean *any* change of mind. There is nothing praiseworthy about arbitrarily adopting and dropping different views. By "change of mind," we mean changing one's beliefs in response to the evidence.[3]

But despite praise for individuals' willingness to change their minds being a familiar practice, it is worth taking a step back to consider just how peculiar this practice is. After all, in changing our minds in response to the evidence, it seems that we are simply acting as we ought to epistemically. We *ought* to adjust our beliefs in response to the evidence. Indeed, if anything is obligatory for us epistemically, then certainly it is these adjustments! Why all this fuss, then, about simply doing what we ought – about meeting a minimum epistemic standard?

A tempting answer is that we praise figures like Putnam because they seem to be one step closer to approximating the epistemic ideal: the figure who gets things right

[1]The reference here to Ravitch's "open-mindedness" is discussed in section 6.

[2]DiPaolo's (2021) examples of reluctant converts are a potential subset of the cases I am interested in, though weighty costs for the conversion are key for my purposes. Schulz (2010) also includes many helpful examples of this phenomenon, such as the cases of Penny Beernsten and C.P. Ellis.

[3]Unless otherwise specified, when I say "change of mind" going forward, I mean in response to the evidence. I further clarify the notions of evidence, responsiveness, and mind changing in section 3. For discussion of changing your mind *without* gaining new evidence, see Woodard (2022).

https://doi.org/10.1017/epi.2024.54 Published online by Cambridge University Press

epistemically straightaway – that is, being the kind of agent who does not commit these errors or engage in these deficient epistemic practices in future. But if we end our story there, then we seem to have lost sight of just how mysterious our praise of these figures is. On this view, the praise in our opening examples simply serves to draw our attention to the actual epistemic ideal. There is nothing particularly special, then, about Putnam or Einstein or any of these cases, despite the tenor of the praise. We praise them because they represent an important stepping stone for realizing the epistemic ideal.

The very notion of an "epistemic ideal" will be in question in this discussion, so we should be more precise. Call the epistemic ideal at stake here the "Traditional Ideal": the ideal epistemic agent responds to the evidence just as they ought, that is, without any errors or deficient epistemic practices.[4] To give a name to the corresponding view, call it the "Traditional Ideal View": we praise agents who change their minds in light of the evidence because they are closer to realizing the Traditional Ideal.

It might seem that the Traditional Ideal View captures all we should say about these cases. There is of course nothing infallible in our practices of praise. If there is concern about the level of enthusiasm behind this praise, a defender of this approach might say that these paeans are simply out of proportion to their target and can be moderated accordingly.

But, in my view, we are not just highlighting a step in a path to realizing the Traditional Ideal in these cases. There is something uniquely praiseworthy that we are tracking and for which the pitch of this praise is fitting. What we learn about agents in these cases, I argue, is that they are *committed* to epistemic goods, such as knowledge, truth, understanding, or justification. This commitment to epistemic goods is, furthermore, not something that we can be confident obtains even in the case of the Traditional Ideal. I argue that it is compatible with being epistemically ideal in this sense – with responding to the evidence without the diversion of any error – that one is *not* committed to epistemic goods. Agents who change their mind in response to the evidence and, importantly (as I will draw out), at a likely cost to themselves give us the best indication we can have that they have a genuine commitment to epistemic goods.

In the paper's second section, I explain the stakes of this discussion. I argue that we face the pervasive threat of what I call the "Problem of Mere Epistemic Performance" – uncertainty regarding whether agents, including ourselves, have a commitment to epistemic goods or whether they are merely performing this commitment in order to serve their non-epistemic ends. In the third section, I clarify my view of commitment and other terminology. In the fourth section, I show that the Traditional Ideal is compatible with lacking a commitment to epistemic goods and therefore should be rejected as the epistemic ideal. In the fifth section, I argue for the surprising conclusion that the ideal epistemic agent *must* have certain false beliefs or deficient epistemic practices. These failures are necessary for them to possess a genuine commitment to epistemic goods. In the sixth section, I respond to various objections.

2. The Problem of Mere Epistemic Performance

We live in a world awash with merchants of doubt, bad faith actors, shills, the rapacious, and the megalomaniacal. Alexander Guerrero argues that the pervasiveness of these figures and contexts where "the experts in question may have or appear to have interests – economic, professional, personal – that are in less than complete alignment with the non-experts" poses a serious threat to our epistemic dependence – that is, the necessity of our trusting experts to gain the vast majority of our knowledge about the world (2017: 157). Because we know such figures abound, we often have good reasons "not

[4]The Traditional Ideal is discussed in more detail in section 4.

https://doi.org/10.1017/epi.2024.54 Published online by Cambridge University Press

to completely trust the expert(s)" (157). Part of what makes these contexts so difficult to navigate is that agents who have no investment in epistemic goods but push whatever line best serves their non-epistemic interests are very skilled at *performing* this investment: they do not shy away from – they even relish – opportunities to present and argue for their views because this performance generates strategically important credibility and can push others to adopt positions advantageous for their non-epistemic interests.[5]

Suppose, for example, I am a public relations expert tasked with defending a corporation's actions. I do not present myself as only offering arguments that serve the corporation's agenda above all else and that I do not actually believe. I instead present myself as offering good arguments and evidence that have nothing to do with who pays my salary and everything to do with their merit. But, in fact, everything I argue *is* driven by a concern with protecting the corporation's interests, not epistemic goods. Or suppose I am a researcher hired by the same corporation and my position depends on producing research in line with the corporation's interests, so that's what I do. But this is not how I present my work publicly; I instead present it as a genuine pursuit of the truth. We also do not need to paint with a Machiavellian brush. Such figures may be entirely unaware that they are beholden to a set of non-epistemic ends that are guiding their activity, and this guidance may be much more insidious. Orders and threats are not always explicit.

Expanding to more ordinary contexts, motivated cognition – where "evidence is sampled and critiqued selectively in order to reinforce what one *wants* to believe" – seems to be ubiquitous in shaping our interactions, perhaps especially when we form political beliefs (Hornsey and Fielding 2017: 460). Someone engaged in motivated cognition presents themselves as invested in the truth and epistemic merit of what they believe, but are in fact driven primarily by a desire to engage in – among other behaviors – identity signaling and protection, adopting beliefs that flatter and support their in-group and stigmatize out-groups. When they encounter counterevidence that conflicts with their political interests, they dismiss it, and when they encounter evidence that conflicts with their political opponents' interests, they play it up. But this is of course not how they present or view themselves: they take themselves to be invested in the evidence for and truth of their beliefs and try to engage others in giving and asking for reasons, even though they have no ultimate investment in the fruits of this practice.[6] Anyone who has attempted to engage politically with others online will be very familiar with this phenomenon: the superficial appearance of giving and asking for reasons can easily be an attempt to bully others or a way of broadcasting one's political allegiances.

Of course, it is not just other people who are the source of vitiated epistemic practices born of ulterior motives. We ourselves can be subject to all of these same forces, warping our deliberation, convincing us that we are after the truth when in fact we are simply looking to endorse whatever serves our interests.[7] I might wonder, for example, whether I really believe that our nation's war is a noble cause, as I claim, or whether this belief is the product of a reflexive nationalism I have been raised with. Nietzsche famously argues that philosophers are particularly prone to this self-deception: we envision ourselves as driven by a will to truth but our arguments are better read as "the personal confession of its author and a kind of involuntary and unconscious memoir" aiming to impose our preferred, implicit morality, rather than a product of following the evidence where it leads (1886/1989, 13).

[5]I agree with Briana Toole's observation that there is no straightforward way to draw the "epistemic" versus "non-epistemic" distinction, but "paradigm examples give us a strong grasp on the category of epistemic factors: the examples are evidence and justification, truth, belief, reliability – in short, those properties that featured in the post-Gettier trend that characterizes traditional epistemology" (2022: 49).

[6]For a philosophically informed overview of the motivated cognition literature, see Bardon (2020).

[7]On this type of reflexive case in the context of motivated cognition, see Carter and McKenna (2020).

https://doi.org/10.1017/epi.2024.54 Published online by Cambridge University Press

In all of these cases – from the compromised "experts" to motivated cognizers, including ourselves – what we find are agents who are going through the motions of an investment in epistemic goods, but who are ultimately driven by their ulterior, non-epistemic interests. When we encounter, then, the trappings of genuine epistemic activity but also have reason to believe that this activity serves the agent's non-epistemic interests, we often become skeptical of the arguments, reasons, and evidence being advanced in order to protect ourselves from being misled or deceived. Call this the "Problem of Mere Epistemic Performance" – uncertainty regarding whether agents, including ourselves, have a genuine commitment to epistemic goods or whether they are merely performing this commitment in order to serve their non-epistemic ends.[8]

A great deal rides on this assessment. Our responses to figures who have and do not have this commitment should be, in many cases, radically different. If my interlocutor is committed to epistemic goods, then engaging them with arguments, normative reasons, and evidence is fitting. This is true even if they are, epistemically, in disastrous shape – even if they subscribe to a host of egregiously false beliefs and terrible epistemic practices. If they are nonetheless committed to epistemic goods, then engaging them with arguments, reasons, and evidence is fitting because they take themselves to be accountable to what is right epistemically and are in principle responsive to these considerations, even if getting them to appreciate their force will take enormous effort and has no guarantee of success.

But if my interlocutor is *not* committed to epistemic goods, then engaging them as if they were is often a grave error.[9] Because the agent lacks this commitment, they will not ultimately be moved by arguments, reasons, or evidence that conflict with their non-epistemic ends. It is therefore often a waste of time and effort to try and engage them on these scores – time and resources that could be much better spent on other pursuits. For example, if we object to the non-epistemic ends and interests this agent is serving – say, the political projects or institutions they are beholden to – we would be far better off working to combat those interests directly and doing so in ways most strategically appropriate for our goals – strategies that will likely *not* involve trying to engage actors tied to these interests in the pursuit of inquiry. We generally should not spend our time, for instance, attempting to persuade the public relations expert or the researcher bound by these institutions even though they present themselves as open to such persuasion – for example, attempting to refute every one of their arguments (of which they perpetually produce more) or presenting them with yet more evidence when they claim they are skeptical of anything we have presented previously. Rather than attempting joint inquiry, we should instead focus our efforts on combatting their deception by, say, advocating for policy that curbs the influence of these actors on legislation.

Similarly, once I judge that an agent is engaged in motivated cognition despite presenting themselves as truth seeking, then the nature of fitting engagement shifts. The appropriate response does not seem to be "explication" or presenting "the evidence in language that is as clear and digestible as possible" (Hornsey and Fielding 2017: 459). Instead, we need to locate the relevant "'attitude roots': the beliefs, ideologies, fears, and identity issues that motivate people to want to reject" inconvenient truths, and then in order "[t]o effectively convert people . . . tailor communication strategies to work with – rather than against – these underlying motives" (460).[10]

[8]See Kornblith (1999) and McBrayer (2024) for related discussion.

[9]See McKenna (2023: Chapter 6) for related discussion.

[10]Elsewhere (e.g., Shields (2022)), I question the type of conceptual framework Hornsey and Fielding use for their specific discussion of "conspiratorial ideation," but here, I am concerned with their broader overview of the motivated cognition literature.

https://doi.org/10.1017/epi.2024.54 Published online by Cambridge University Press

In responsibilist virtue epistemology, a version of the Problem of Mere Epistemic Performance looms large. Linda Zagzebski has argued that "[t]he simplest way to describe the motivation of the intellectual virtues is to say that they are all based in the motivation for knowledge. They are all forms of the motivation to have cognitive contact with reality" (1996: 167). This "motivational component . . . must be inculcated sufficiently to reliably withstand the influence of contrary motivations" (178). But we might wonder how we can be confident that an agent in fact has this overriding motivation for knowledge. Jason Baehr similarly argues that "an intellectual virtue is a character trait that contributes to its possessor's personal intellectual worth on account of its involving a positive psychological orientation toward epistemic goods" (2011: 102).[11] Whether, for example, an action is an "instance of open-mindedness depends in part on the immediate *motivation* behind it. Imagine a person who sets aside or moves beyond one of his beliefs in order to assess an argument against this belief, but who has no real intention of making an honest or fair assessment of this argument (he just wants to get the attention of his interlocutor, say). Intuitively, this person is not genuinely open-minded" (151). But despite the crucial role that this positive orientation to epistemic goods plays in determining whether a given action counts as intellectually virtuous, Baehr and Zagzebski do not tell us how we can be confident that an agent actually possesses this orientation. Baehr's own examples of the mixed motives behind agents' actions show how difficult it can be to make these judgment calls.

3. Commitment to epistemic goods and other terminological clarifications

We are searching, then, for a way of determining whether the agents we engage with, including ourselves, in fact, have a commitment to epistemic goods. But what do I mean by "commitment"? Here I follow Cheshire Calhoun's account. For Calhoun, "a commitment is both an intention to engage with something (a person, relationship, goal, activity, identity, etc.) and a preparedness to see to it that that intention to engage persists," even when various temptations to abandon the intention arise or one's desires change (2009: 618). Crucially, for Calhoun, "[a]ll genuine commitments are active in the sense that they are made, not merely discovered as facts about one's psychology, and they persist through being sustained" (617). On this view, the existence of a commitment is necessarily bound up with one's actions. We do not have a commitment simply because we say we do and even if it appears phenomenologically obvious that this commitment obtains.

To adapt one of Calhoun's examples, suppose I take myself to be committed to the study of biology. But as soon as I fail to get an A in a class, I stop pursuing the subject. In this case, I was convinced that I had a commitment. But if I did, then I would not have been deterred by the very first disappointing outcome. On reflection, it becomes apparent that I have been driven by retaining a flattering self-conception (as an A student), rather than a commitment to study the subject. Similarly, suppose I take myself to be committed to a certain political project. But as soon as circumstances call for acting on that commitment in a way that would be inconvenient or costly for me, I decide not to take action, and whenever this conflict subsequently arises, I also do not take action. In that case, it seems clear that I lack this commitment despite my previous self-conception. As Calhoun puts it, "We measure depth of commitment by what a person is prepared to do or to resist in order to see to it that the intention to engage persists" (618).

Conversely, I might discover that I had a commitment I didn't realize I had. Consider the action movie trope of the seemingly callous, self-absorbed mercenary hired in the service of a righteous cause. When a conflict subsequently arises between benefiting themselves and

[11]Italics removed to make the passage easier to read.

https://doi.org/10.1017/epi.2024.54 Published online by Cambridge University Press

abandoning those they were hired to protect, they opt – to their own surprise – to support the cause even at a substantial cost to themselves. It turns out their self-conception was misguided: they were not simply driven by greed above all; they have a commitment to doing the right thing. Ruth Chang similarly emphasizes that although commitment is active in that it needs to be sustained in the face of incentives to abandon it, it need not be the product of a conscious decision: a commitment can result from "an unconscious and non-deliberate decision; after living together for a few years, more and more of your long-term plans involve Harry, and his interests have greater importance than they had before. Indeed, were he to need a kidney, you would offer up one of yours. You have resolved [to undertake a commitment to him] . . . but not consciously or deliberately" (2013: 85).

This view of commitment where "genuine commitments are made, not merely discovered as facts about one's psychology," also implies a certain overlap between the metaphysics and epistemology of commitment. A commitment exists only insofar as we take certain actions on this view. But actions unfold over time, often in complex, open-ended ways.[12] In many cases, we cannot just read off isolated time slices exactly what an agent is doing, even when that agent is ourselves. Am I, for example, training for a marathon or just doing daily runs? Am I writing a book or merely jotting down some notes? Am I apologizing to a friend or simply placating them? Answers may not be apparent, even to ourselves, for some time.

Given, then, that certain actions are necessary to count as having a commitment and that the performance of an action is not always a settled datum, it will sometimes be more apt to say that an agent's activity or actions provide *evidence* of a commitment – a formulation I sometimes take advantage of in this discussion since it helpfully conveys that this judgment may not be dispositive. Alternatively, it will also sometimes be more apt to say that the action or actions *establish* the existence of a commitment – a formulation I also use – because taking these actions can suffice for having the commitment on this view. In my view, this is a feature, not a bug, of the account: it correctly reflects both the deep connection between commitment and action as well as the complex dimensions of action.[13]

In this discussion, I am concerned with the commitment agents may have to epistemic goods specifically. We saw in the previous section why determining the existence of this commitment is crucial: it helps us navigate the Problem of Mere Epistemic Performance. Given this account of commitment, then, whether an agent has a commitment to epistemic goods will depend on what they are "prepared to do or to resist in order to" do what is right epistemically. We will therefore need to observe the costs that the agent is willing to bear in order to determine whether this commitment exists. Without experiencing these costs and persisting in one's actions despite them, it is simply unclear whether any such commitment obtains.

I have also referred to agents changing their mind in response to the evidence, but I should clarify that I intend this to be read as *good* evidence. What makes evidence good is whether that evidence increases the justification or reasonableness of the belief. There are of course difficult questions about what this relation looks like and how we assess it, but for the purposes of this discussion, we only need to assume *that* there is such a thing as good evidence and that agents can be exposed to it.

[12]See Tenenbaum (2021) for recent discussion.

[13]Rejecting this tight connection between commitment and certain actions requires defending the view that an agent can possess a commitment even if they endure no costs. The fallibility of our phenomenology the examples in this section demonstrate and the apparent centrality to commitment of its persistence *despite* obstacles, temptations, or costs make defense of this view an uphill battle.

https://doi.org/10.1017/epi.2024.54 Published online by Cambridge University Press

Let me also say more about what I mean by "changing one's mind." I specifically have in mind cases where agents change their beliefs by acknowledging an error on their part: that their beliefs turned out to be false, that they did not have the justification for their beliefs that they thought they had, that they did not fully understand a crucial point, that they made a series of misguided inferences, etc. – changes that we see reflected in the opening cases.

Not every type of epistemic error will be present in all cases. For example, I might go from having a false, but entirely reasonable belief in response to an initial body of evidence to having an equally reasonable, but now true belief. Many scientific cases, it seems, will have this shape. Alternatively, I might go from having a true belief where I lack understanding to what turns out to be a false belief where I nonetheless gain a great deal of understanding.[14] Furthermore, different types of errors and different levels of cost agents face will likely also affect how much praise we think they are due for their change in view. I am most interested in cases such as the opening ones where the mistakes are clearly acknowledged and multiple (errors regarding what the agent took to be true, errors in their original inferences and understanding, etc.), the costs to the change are high, and the praise in response is substantial.

By "costs," I have in mind material costs (loss of income, jobs, etc.), social costs (loss of friendships, communal ties, etc.), or personal costs (an injury to one's pride, loss of personal satisfaction, etc.). There are also substantial costs, that is, costs that are not simply outweighed by benefits the agent might acquire by changing their mind. For example, I might experience temporary embarrassment at having to admit that I was wrong, but if I stand to accrue massive financial benefits from changing my mind, then this is a good reason to think I am not incurring serious costs. It can be difficult to say what the costs to a specific agent are of a particular decision, and anticipated, substantial costs that we have every reason to think will be suffered of course may not end up materializing due to exceptional or interfering circumstances.[15] But *that* there are likely costs and that they can be more or less weighty, I assume in the following.

I also say that agents change their minds in response to the evidence, but what do I mean by "in response to"? The phrase evokes the epistemic basing relation, that is, that agents will change their mind *on the basis* of the good evidence they are encountering (Korcz 2021). But I opt for the more neutral formulation of "in response to" because the crucial question at stake in evaluating these cases is whether the agents are indeed basing their beliefs on the good epistemic reasons they have or on some other non-epistemic motive. Another way to construe my overall position is that changing one's mind in response to good evidence and at a likely cost to oneself is the best indication we have that the agent *is* basing their beliefs on a commitment to epistemic goods rather than non-epistemic interests. The more neutral formulation of "in response to" avoids baking into these cases that the commitment already obtains.

4. The know-it-all

It is not uncommon in everyday life and in fiction to encounter a figure who uses what they claim is their superior knowledge as a cudgel to one-up, intimidate, or even tyrannize over others – sometimes they practice their art effectively, sometimes not. We

[14]This diversity in epistemic mistakes is important because it leaves room for a pluralism about epistemic goods. If this increasingly popular pluralism among epistemologists is right, then we do not want to insist that changing from false to true beliefs is all that matters. I discuss this pluralism in more detail in the next section. Thank you to an anonymous reviewer for pushing me to say more on this point.

[15]Hence the qualifier "likely" in front of "costs," which should be assumed throughout unless otherwise indicated.

https://doi.org/10.1017/epi.2024.54 Published online by Cambridge University Press

have a name for these figures. We call them "know-it-alls." But the point, of course, is that despite their self-conception, the know-it-all *doesn't* know it all. They are self-deceived.

Suppose, however, we have a figure who *really* does know it all. More specifically, consider a figure who is the direct embodiment of what many philosophers have taken to be the goal of our epistemic lives – namely, acquiring positive epistemic states, such as true beliefs, justification, or understanding, and no negative ones, such as false beliefs, lack of justification, or lack of understanding. Whenever this figure encounters some question, proposition, or piece of evidence, they have exactly the right epistemic relationship to it. Such a figure would be an embodiment of what I called earlier the Traditional Ideal. In the latter half of the twentieth century, "the epistemic goal [was] . . . standardly taken in epistemology . . . to be that of getting to the truth and avoiding error" (Kvanvig 2005: 285). This monistic veritism was then challenged by increasingly pluralist views of the epistemic good: "[I]t is intuitively plausible that the intellectually virtuous inquirer not only loves truth for its own sake, but loves a host of other epistemic goods for their own sake as well . . . Examples of such goods would be knowledge, understanding, wisdom, and reasonable belief" (Greco 2021: 53). But whether one is a veritist or pluralist about epistemic goods, the point is that the ideal epistemic agent always acquires the positive epistemic state and not the negative one. For a proponent of the Traditional Ideal View, the agent who changes their mind is therefore a step closer to realizing this ideal: they are a step closer to acquiring the positive epistemic state in the future *right away*, that is, without any detour into an error that previously led to a negative epistemic state and therefore generated the need to change their mind.

In a moment, I want to explore a very specific version of the Traditional Ideal. But it is worth first dwelling on another question that will become central to this discussion: the satisfaction we can derive in general from getting things right.[16] There are of course all kinds of epistemic satisfaction: we might be proud of ourselves for recalling a hard-to-remember fact, finding a flaw in an argument, or learning a new idea. But there are also less innocent forms of satisfaction in these contexts. We might, for example, take satisfaction in these very same phenomena – recalling a fact, finding a flaw in an argument, learning a new idea – because we assume we will be admired or reap certain rewards or because we will be able to use these phenomena to embarrass those we take to be our adversaries. Simply consider the characters canvased at the beginning of section 2.

This type of satisfaction might also be more general. We might just derive a deep pleasure from being right epistemically. When we are right about the facts and our interlocutors wrong, and if the surrounding dispute is heated (say, a political dispute where the epistemic question at stake is whether something is a fact), it is hard not to feel a certain satisfaction in being proved right.[17]

Let's combine these observations with our Traditional Ideal. When we do, we end up with the following figure:

> *The Self-Serving Know-It-All*: They respond to every epistemic phenomenon just as they ought.[18] That is, their beliefs are always true, and they always have the best justification for their beliefs. They weigh evidence exactly as they should, and they

[16]See Roberts and Wood (2007: 153–4 and 173–7) for related discussion and analysis of Aristotle's and Locke's views on epistemic satisfaction.

[17]Chapter 1 of Schulz (2010) provides various delightful examples of epistemic self-satisfaction.

[18]A central argument of this paper is that the Self-Serving Know-It-All does *not* in fact respond to every phenomenon just as they ought – that this figure is not our epistemic ideal – so the "ought" here should not be read as settled, but open to contestation.

https://doi.org/10.1017/epi.2024.54 Published online by Cambridge University Press

> appear to evince all of the intellectual virtues. But, as it turns out, this figure aims to be – and are – epistemically impeccable because they derive a self-serving satisfaction from achieving this state; they find it deeply satisfying to be right epistemically. Their self-interest happens to align perfectly with ideal epistemic conduct.

By "self-interest" here, I have in mind a straightforward version of what Robert Shaver calls a preference or desire account that identifies "self-interest with the satisfaction of one's desires. Often, and most plausibly, these desires are restricted. What makes a desire self-regarding is controversial, but there are clear cases and counter-cases: a desire for my own pleasure is self-regarding, a desire for the welfare of others is not" (Shaver 2023). In what sense, then, is the self-interest of our know-it-all served by being right epistemically? The answer is that we can substitute any of the candidates from above. For example, they may desire this rightness because they simply enjoy the experience of being right epistemically. Or they may desire it because they think it will always further their non-epistemic projects. Or, to adopt a more Nietzschean diagnosis, they may desire it because they are convinced that having the optimal epistemic status is the best means of ultimately imposing their will on others and the world – among of course many other possibilities.

Now what is wrong with the Self-Serving Know-It-All exactly? In my view, the problem with this figure is that they have only a *contingent* relationship to epistemic goods: they do not value epistemic goods for their own sake, but because it aligns with their self-regarding ends. If, then, the world were in a state where this figure judged their own satisfaction to be better served by deficient epistemic conduct, then they would engage in that deficient epistemic conduct. As it turns out, their actual world *is* in a state where they judge that their satisfaction is best served by this ideal epistemic conduct, and this figure is constituted in such a way that they can seamlessly carry out the corresponding epistemic performance.[19] But this might not have been the case. And I submit that there is something lacking epistemically for the Self-Serving Know-It-All as a result of their contingent relationship to epistemic goods. It would be better, from an epistemic point of view, to have a noncontingent relationship to epistemic goods – to value and be committed to epistemic goods for their own sake. It would be better because it means that if the world *were* different, excellent epistemic conduct would still follow, whereas, for the Self-Serving Know-It-All, this conduct would disappear in very close possible worlds.

By contrast, the opening examples of the agents who change their minds in response to the evidence gives us excellent reason to think that they do have this commitment to epistemic goods. To fully appreciate this point, we need to highlight a further dimension of these cases. It is crucial that there is a likely cost to the agent who changes their mind. In some instances, the cost is material: it may require giving up lucrative career opportunities or one's established income for an uncertain future (along with the many other material risks that attend economic precarity). In other instances, the cost is reputational: the agents are in fields or areas where getting things right epistemically matters a great deal, and they will have to concede that they have failed in this goal, including on weighty questions. Similarly, some onlookers will naturally wonder: if they could be wrong about this question, maybe they could be wrong about other questions. That is, their trustworthiness can be called into question – a devastating result for a professional inquirer. The agent may also face social costs, such as breaking deep ties with a community who shared their earlier views, resulting in damage to, or the end of, long-standing relationships and possible backlash from that community. Conversely, those who share the agent's novel view may remain skeptical that the agent's change is

[19]When I say "judgment" here, this does not mean that these views are arrived at through conscious deliberation. Such judgments can be arrived at via subpersonal mechanisms.

https://doi.org/10.1017/epi.2024.54 Published online by Cambridge University Press

genuine, prompting stigma there as well.[20] These agents nonetheless persist, despite these costs, in judging their previous view as false or deficient.

But why think that the evidence in such cases is *especially* compelling that we are dealing with an agent with a genuine commitment to epistemic goods? This is because it goes as far as we can in ruling out the possibility that the agent only has a contingent relationship to epistemic goods – that even if their self-interest is threatened, and threatened profoundly, they will nonetheless do what is epistemically called for. If their epistemic conduct is impeccable but always aligns with their self-interest, then we simply cannot be confident that they have any such commitment.

My observations here are, in a way, the epistemic counterpart to Kant's views on moral duty and inclination. Kant famously tells us that the shopkeeper who does "not overcharge his inexperienced customer" does the right thing, "but this is not nearly enough for us to believe that the merchant proceeded in this way from duty and principles of honesty; his advantage required it" (1786/2011: 23). But for cases where personal satisfaction *is* at odds with doing our moral duty, we can be confident that the agent has a genuine commitment to what is morally right. In fact, Kant is even more precise here: we should try to isolate cases where individuals cannot gain any satisfaction from their actions, *including* satisfaction from doing what is morally right (in addition to other costs they may incur). He asks us to consider the once compassionate individual who previously derived pleasure from helping others but is now "beclouded by" their "own grief" (25). Were these individuals to "tear" themselves "out of this deadly insensibility" and respond to "the need of others" not because they are "sufficiently occupied with" their own needs but "solely from duty," then their action would have "genuine moral worth" (25). The logic behind my critique of the Self-Serving Know-It-All and my case in favor of the epistemic superiority of the figures who change their minds at a cost to themselves is the same: we can be confident that these figures and not the Self-Serving Know-It-All have a genuine commitment to epistemic goods.[21]

If we wanted to give a label to the broader phenomenon at stake, we could call it "epistemic selflessness." We evince epistemic selflessness whenever we prioritize the epistemic goods over our own interests. Note that the category would encompass a wider class than the case of changing one's mind at a likely cost to oneself. I display epistemic selflessness when, for example, I exhibit intellectual courage, retaining a view I take to be epistemically well-supported in the face of "an apparent threat to one's own well-being" (Baehr 2011: 177). But epistemic selflessness is *most* apparent when I not only face likely material and social costs for my view, but when I cannot even gain any sort of epistemic satisfaction from my view – that is, when my view involves conceding that I was wrong. In cases where an agent appears to evince intellectual courage, by contrast, they might be driven by the satisfaction they derive from taking themselves to be right. They might, for example, derive satisfaction from the image of themselves as a kind of epistemic martyr or renegade. The latter persona has a particular antipathy toward consensus: when they observe (what they take to be) areas of agreement among other intellectuals, they see only homogeneity and groupthink anathema to serious inquiry. They prize above all the image of themselves as one of the only remaining free thinkers in a sea of conformity. Now suppose in a certain case, this figure is actually right about a given issue and the group they define themselves in contrast to is wrong, and their being right did come with professional and material costs. Even so, given the satisfaction they derive from viewing themselves as intellectually courageous, it is not clear how much evidence we get that this figure is committed to epistemic goods. If, for instance, the tables were turned and the

[20]See Tosi and Warmke (2022) for discussion of social costs.

[21]For exegesis on these claims from Kant, see Herman (1981).

https://doi.org/10.1017/epi.2024.54 Published online by Cambridge University Press

group was right and the figure was wrong about the issue, they would be unwilling to acknowledge their error. This is why cases of changing one's mind at a cost to oneself are better evidence of commitment to epistemic goods: they come closest to ruling out the possibility of self-serving motivations since they deny the figure satisfaction even from being right epistemically.

A worry may arise at this point, however, that even if everything I have said is right, all it gives us is a criterion for *demonstrating* to others that an agent is epistemically ideal. But it doesn't follow that the epistemic ideal should include features that simply make one's ideality apparent to others. Suppose, for example, I have a terrible poker face in a moment of courage, revealing my fear to others as I push through. It would not follow, however, that having a terrible poker face makes me *more* courageous or a better ideal of courage.[22]

Two points are worth clarifying here. First, my account is not just intended to give *others* an indication of one's epistemic ideality. As I argue in section 2, the Problem of Mere Epistemic Performance applies just as much to oneself as it does to others: I can discover about myself that I have a genuine commitment to epistemic goods when I change my mind. Second, while having a bad poker face does not necessarily enhance the quality of one's courage, having to overcome an obstacle often indicates a superior instantiation of a virtue. For example, the agent who overcomes fear in order to perform a courageous act is plausibly more courageous than an agent who experiences no fear or friction of any kind in carrying out the act. We might even describe the latter as more of a "knack" rather than a full-fledged virtue, at least in certain cases – an uncritical disposition, rather than a product of a "drive to aspire" and an articulate deliberative process (Annas 2011: Chapter 3). Alternatively, suppose that this frictionless agent has until now never experienced fear when performing a courageous act, but then once they finally do have this experience, they fail to perform the relevant act. To outsiders, it might appear that the case of frictionless courage prior to this point is the paradigm case, but with full information, most will conclude that the agent who had to conquer their fear was more courageous.[23]

More generally, considering an agent's internal life and broader activity is crucial for assessing the nature of an agent's action and character.[24] Learning that an agent had to overcome obstacles to realize a virtue – whether courage, generosity, temperance, etc. – often leads to the judgment that their action is a superior representation of the virtue than an agent who experienced no such resistance. The reasoning behind this judgment echoes the one I have argued for: that an agent will perform the act even when they have

[22]I am grateful to an anonymous reviewer for clearly articulating this worry; I have used their formulation of it here.

[23]This puts my view seemingly at odds with John McDowell's account of the virtuous agent, which says that the virtuous agent does not – as the continent agent does – weigh up reasons for and against acting in a virtuous way, but simply perceives the right thing to do, silencing any reasons against it (1979: 333–335). McDowell's account, however, is vulnerable to worries I outline here – namely, that a virtuous agent who simply "perceives" the right thing plausibly has more of a knack-like relationship to virtue because they do not apprehend the force of competing reasons. This raises worries about the contingency of this figure's conduct, were such competing reasons to eventually become live (though McDowell would presumably deny this possibility). As Annas points out, such a picture also threatens our ability to give "any plausible account of ethical advice or disagreement. A loyal person, say, is asked what someone should do who wants to be loyal to their friend, but has realized that the friend is taking drugs. We expect the person asked for advice to be able to offer reasons for and against breaking off relations with the friend. It would obviously be absurd if they replied that there was no way they could explain; the questioner should just watch some loyal people and pick up what they do" (2011: 26). There are larger questions here about the relationship between ideality and continence worth exploring, but I leave them for future work.

[24]See Kamtekar (2004) on a similar point in her response to situationist critiques of virtue ethics.

https://doi.org/10.1017/epi.2024.54 Published online by Cambridge University Press

strong, felt incentives not to means we can be confident that they genuinely, not just contingently, possess this virtue.

Suppose, then, we grant that changing one's mind at a likely cost to oneself is the best evidence of epistemic selflessness, and therefore, the best evidence we can have that an agent has this commitment to epistemic goods. Why think this is a particularly striking result?

5. The Traditional Ideal rejected

A key implication of my account is that we should abandon the Traditional Ideal. If we agree that it is better to have a noncontingent commitment to epistemic goods, and we agree that the best evidence for this commitment is the willingness of an agent to admit they are or have been mistaken and changed their mind accordingly, then the ideal epistemic agent must have at least *one* false belief or *one* deficient epistemic practice. These failures are necessary because they provide the agent with an opportunity to establish their commitment to epistemic goods. Absent these failures, we simply cannot have the same confidence that this commitment obtains. Why? Because the possibility remains that the agent may be a Self-Serving Know-It-All. Their apparently epistemically impeccable conduct would disappear as soon as they take their self-interest to be threatened by this conduct. It just so happens that these circumstances have not arisen.

If epistemic failures and changes of mind are therefore our best evidence of an agent's commitment to epistemic goods, then we arrive at the very surprising conclusion that an agent with false beliefs or deficient epistemic practices can be epistemically more ideal than an agent with only true beliefs and no deficient epistemic practices. Instead of the Traditional Ideal, then, we find that the epistemically ideal agent must have certain false beliefs and deficient practices in order to establish a genuine commitment to epistemic goods.

To draw out just how surprising this conclusion is, consider Alex Worsnip's recent discussion of agents who hold what he calls "suspiciously convenient beliefs," that is, beliefs that align with the agent's interests and desires (2023). According to Worsnip, we always have *pro tanto* higher-order evidence that suspiciously convenient beliefs are irrationally held because such beliefs are plausibly formed as a result of motivated cognition rather than a genuine concern with the truth: I believe p because it serves my interest and desires, not because of its truth. A natural reply to Worsnip's argument is to simply deny the significance of this higher-order evidence: the evidence in favor of a first-order proposition p settles the question of whether to believe p; higher-order evidence concerning one's desires or interests for believing p is therefore irrelevant. Worsnip replies that such a view would only work for a "truly ideal epistemic agent [who] would always be able to look at the first-order arguments and evidence bearing on their belief, and infallibly assess whether they support the belief in question" (253). In other words, the ideal epistemic agent will always have the right orientation toward the first-order evidence, so they do not need to consult higher-order considerations about their motivations. But Worsnip points out that we are *non*-ideal agents, so higher-order evidence *is* relevant to our deliberations and should lead us to downgrade our confidence in p when p would serve our interests or desires, thereby indicating the possibility of motivated cognition. In Worsnip's view, it follows that repudiating the relevance of higher-order evidence may well be legitimate in ideal epistemology, but it is a serious mistake for those pursuing non-ideal theorizing relevant to the kinds of creatures we actually are.

Despite this critique of ideal epistemology and its examination of the role of self-interest in warping deliberation, Worsnip's account leaves intact the Traditional Ideal. One might even use it to push back against my figure of the Self-Serving Know-It-All. Worsnip agrees with the ideal epistemologist that given that the ideal epistemic agent's

https://doi.org/10.1017/epi.2024.54 Published online by Cambridge University Press

"determinations of what the first-order evidence supports are always correct, any evidence suggesting that these determinations are incorrect is *ipso facto* misleading, and thereby liable to lead them astray" (254). Worsnip is therefore aligned with the ideal epistemologist in that both endorse the view that the agent with flawless epistemic judgment is not lacking in any way. And as obvious as this view might seem, my discussion has tried to show why this is misguided.

Where, then, does Worsnip go wrong? Within the terms of my account, he does not consider the possibility that the agent's suspiciously convenient beliefs may not just enhance their external practical "interests and desires," such as their political or financial concerns (to borrow from Worsnip's opening cases); he misses that the agent's interests and desires may also be served *by their very belief that p and when it is the case that p* (239). That is, Worsnip does not consider the epistemic self-satisfaction agents can derive from their successful epistemic practices – from getting things right. But as Kathryn Schulz points out: "[T]he thrill of being right is undeniable, universal, and (perhaps most oddly) almost entirely undiscriminating . . . [W]e can relish being right about almost anything" (2010, 3). Recall my examples of this phenomenon of epistemic self-satisfaction in section 4. Even Worsnip's version of the Traditional Ideal, who has exactly the right assessment of all first-order evidence, may be doing so simply because they derive self-satisfaction from their correct epistemic assessments. But I have argued that such a figure is lacking a commitment to epistemic goods, rendering their apparently virtuous epistemic practices utterly contingent and their ideal status in doubt.

Not only do my arguments therefore show, *contra* Worsnip, that higher-order evidence regarding an agent's motivations *is* applicable even to the agent who always has the right assessment of the first-order evidence (i.e., relevant to the Traditional Ideal), they also show that the very preservation of the Traditional Ideal is a mistake – that we should go much further in our critique of ideal epistemology. The problem is not only ideal epistemology's lack of applicability to real-world, non-ideal agents, as Worsnip argues, but the very Traditional Ideal itself – which Worsnip's critique leaves unscathed, continuing to maintain that a "perfectly rational agent, ipso facto, has no imperfect rationality to take account of" (254). But, as we have just seen, this agent *is* lacking. They lack the best evidence we can have that they are committed to epistemic goods: they have never needed to acknowledge false beliefs or flawed epistemic practices that they have subsequently corrected and done so at a cost to themselves, including being deprived of self-satisfaction. If, as I have argued, such errors are necessary to evince a commitment to epistemic goods, then the Traditional Ideal must be radically revised: the ideal epistemic agent will have at least one false belief or flawed epistemic practice that they have corrected in order to establish their commitment to epistemic goods. I turn now to various objections to my view.

6. Objections

Objection 1

It might be argued that we can preserve the Traditional Ideal and avoid this counterintuitive result that the ideal epistemic agent must have at least one false belief or at least one deficient practice by revising the Traditional Ideal. According to this novel version of the Traditional Ideal View, the ideal epistemic agent not only aims to acquire only positive epistemic states and no negative ones but also has a commitment to epistemic goods. In other words, we can simply add this commitment to the Traditional Ideal.

My reply

Note first that this would not be a minor alteration to the Traditional Ideal. This would involve the addition of a unique condition that one does not regularly find in

https://doi.org/10.1017/epi.2024.54 Published online by Cambridge University Press

discussions of or references to the epistemic ideal outside of responsibilist virtue epistemology. So this would already be a substantial result.

But this attempted revision to the Traditional Ideal nonetheless rests on a fundamental mistake regarding the metaphysics of commitment. Earlier, I endorsed Calhoun's claim that "[a]ll genuine commitments are . . . not merely discovered as facts about one's psychology," but exist insofar as they are "sustained," where this involves taking action in line with the commitment, especially overcoming likely costs to one oneself. We cannot, then, simply insist that an agent has a commitment without any other information about the agent's actions. Commitment does not correspond to a set of "facts about one's psychology." Absent encountering costs to holding the commitment and persisting, it is simply unclear whether this commitment obtains. There is, then, no clear evidence for the existence of a commitment to epistemic goods for the figure of the Traditional Ideal who has only positive epistemic states, no negative ones, and who insists on their commitment to epistemic goods. They need certain false beliefs or deficient practices that they are willing to correct in order to establish that they will prioritize epistemic goods over their own satisfaction.

Objection 2

What if the agent is an environment where changing their mind *is* rewarded? Then it would follow that changing your mind isn't helpful evidence for determining whether someone is committed to epistemic goods. Furthermore, we might think this is what is happening in at least some of the opening cases. Nussbaum's very praise of Putnam, for example, suggests that academic philosophy *will* reward these changes of mind. Others might observe the praise heaped on figures such as Ravitch and conclude that they will likely benefit from engaging in this practice.

My reply

It is not clear to me that there is *any* environment where agents are consistently rewarded for changing their mind. Putnam notes that he is frequently criticized for this practice: "The fact that I change my mind in philosophy has been viewed as a character defect" (1988: xi). I invite the reader to consider other examples of intellectual figures changing their minds. My guess is that they will agree with Nussbaum's observation about philosophers for any intellectual domain: homage is paid to the practice, but examples are rare. But the dearth of cases is also not surprising. In intellectual life, getting things right epistemically is treated as the exclusive goal, and getting things wrong seems clearly at odds with this goal (though of course I am arguing against that view here). Furthermore, conceding that one is wrong invites the charge that one is fickle and untrustworthy on epistemic questions. There are no clear rewards for the practice of changing one's mind in current intellectual life, but there are very clear costs. It is not surprising, then, that we find few exemplars, and it does not seem to be common ground that one is rewarded for changing one's mind. I think this is something we should change. If I am right, then this practice is of paramount importance: it shows that we are genuinely engaged in inquiry. Exactly what form practically integrating rewards for changes of mind might take is beyond the scope of this discussion, but worth thinking about.

The objector might argue, however, that if my proposal were taken on board, then it *would* become common ground that changes of mind are praised and rewarded. In that case, we would eliminate precisely what is supposed to be valuable about this practice – namely, that those who opt for it have to do so at a likely cost to themselves. Engaging in the practice would no longer be costly since agents can be confident of praise and rewards.

But what is this world exactly? It is one where epistemic praise is appropriately keyed to our epistemic practices. If so, then *all* of our praiseworthy epistemic practices will in

https://doi.org/10.1017/epi.2024.54 Published online by Cambridge University Press

fact be praised in this world. But then our best evidence for whether someone is committed to epistemic goods will have to lie in other costs they incur since all good epistemic practices are rewarded. As such, the best evidence for this commitment will still come from those who change their minds because, in doing so, they will have to concede that they were wrong and will not be able to derive the satisfaction from being right straightaway. By contrast, an agent who never has to change their minds because they are always right will be one we cannot be as confident is committed to epistemic goods. We cannot rule out the possibility, even in this world, that their epistemic practices result from a self-interested desire that gives them a contingent relationship to epistemic goods.

The objector might then say that the world is one where *only* the practice of changing one's mind is consistently and appropriately praised, whereas other practices are not. But then I've lost the objection's thread. I can readily concede that in this distant world, the practice of changing one's mind would not be the best evidence of one's commitment to the epistemic good. We would have to consider other epistemic practices in this world to determine if some might be able to offer us better evidence regarding agents' relationship to epistemic goods.

Objection 3

Even if our community does not reward this practice of changing my mind, what if I just happen to be the kind the subject who gets a deep, visceral satisfaction from changing my mind? In other words, what if we reverse the case of the Self-Serving Know-It-All? We would then have a sort of epistemic masochist: they derive pleasure from having to change their mind in light of the evidence. Certainly, this figure is just as contingently linked to good epistemic practice as the Self-Serving Know-It-All.

My reply

It is true that I cannot rule out the possibility of this figure. But then I am not arguing that changing one's mind at a likely cost to oneself is infallible evidence in favor of an agent having a commitment to epistemic goods. I am arguing that it is the best evidence we can have.

We can also observe other aspects of the agent's behavior to make a judgment. The epistemic masochist will be inclined to make all sorts of *obviously* false claims and *obviously* unjustified claims in order to generate pleasure from being corrected. This type of behavior is itself sufficient to cast doubt on whether the agent is committed to epistemic goods rather than some other non-epistemic end.

The objector might reply by modifying their case. Perhaps the figure is not an epistemic masochist, but someone who just derives a deep sense of satisfaction from finding out they are wrong on significant topics. In fact, we might cite Nussbaum on Putnam in support of the possibility of this figure: "[B]eing led to change was to [Putnam] not distressing but profoundly delightful, evidence that he was humble enough to be worthy of his own rationality" (Nussbaum 2016). Suppose Nussbaum is right about Putnam. Then it seems we have a figure who may just be interested in deriving pleasure from changing their mind and would therefore abandon a commitment to epistemic goods in nearby possible worlds if adopting this commitment no longer generated this pleasure.

There are several problems with this line of thought. First, and to reiterate, I accept that it is possible that a figure who changes the mind in response to the evidence and at apparent costs to themselves may nonetheless lack a commitment to epistemic goods. This possibility cannot be ruled out. This practice gives us the best, not decisive, evidence in favor of the existence of this commitment. But the figure who derives a self-regarding

https://doi.org/10.1017/epi.2024.54 Published online by Cambridge University Press

pleasure not from changing their mind in general (i.e., the epistemic masochist) but only on substantive questions is strikingly idiosyncratic. The Self-Serving Know-It-All, by contrast, is simply a natural extension of the self-regarding satisfaction we know that virtually all of us derive from getting things right. That we may not be able to rule out this curious figure is, then, not a cause for concern. I also read Nussbaum not as saying that Putnam chased this "delight" and therefore changed his mind, but that when he did change his mind, this was a happy result because it meant he was indeed committed to epistemic goods (and presumably not driven by "personal vanity," as Putnam says in his praise of Carnap). Third, even supposing Putnam (or a similar figure) did seek out pleasure from changing his mind, then, in a way, Putnam is only a very good, but not ideal example of someone who changes their mind at a likely cost to themselves. We should pursue more Kantian exemplars. It would be better if Putnam *were* distressed by having to change his mind. In this sense, examples such as Bolnick may be more ideal since he describes himself as being "horrified" and "mortified" when it came to discovering and deciding to admit his error (Thompson 2018). In this case, we are even closer to ruling out the possibility that there is any self-serving motivation behind the practice.

The objector might, however, point to a related, but different genre of case to press their skepticism.[25] For example, Joshua Diapolo asks us to consider "a truly open-minded agnostic" who eventually comes to believe in God's existence after genuine, rigorous inquiry (forthcoming: 20). The convert then "becomes famous for having undergone a truly open-minded conversion" and prides himself on this image (20). This resulting pride and fame, however, makes him increasingly dismissive and unresponsive to a legitimate critique of his views because he is convinced of his intellectual virtuousness given his original conversion. This kind of case seems to show that we *can* derive a great deal of pleasure from changing our minds, challenging whether doing so can suffice to establish a commitment to epistemic goods.

There are at least two features of this case, however, that differ from the paradigm cases I am most concerned with. The first is that DiPaolo's convert only engages in a single change of mind and ceases to make subsequent changes even when presented with good evidence. What this suggests is that shortly after the initial conversion, we did indeed have good reason to view this figure as committed to epistemic goods. But without any further changes in the face of good evidence, there is increasing reason over time to question whether this commitment still obtains. This is why Rovelli describes Einstein's willingness to "change [his] ideas, not once but repeatedly" as part of what makes him such an admirable epistemic figure. Similar claims are made about Putnam. These figures do not engage in the epistemic self-licensing of DiPaolo's convert, patting themselves on the back for their initial open-mindedness and then refusing to entertain any other revisions; they instead continue to revisit and change their views despite the costs of doing so.

A second difference is that DiPaolo's convert ends up acquiring a great deal of rewards for the change in view, and these rewards play a key role in corrupting his epistemic behavior. This is why the cost constraint is crucial to the cases I am interested in: the less likely the costs and the more likely the rewards for a change of mind, evidence of a genuine commitment to epistemic goods increasingly fades.

Objection 4

It might be argued that the Traditional Ideal preserves certain epistemic perfections that my revised ideal does not – namely, not making any epistemic missteps – and should therefore still be preferred, at least in certain contexts. Perhaps, for example, we

[25]An anonymous reviewer helpfully pointed to the relevance of this "zeal of the convert"-style case.

https://doi.org/10.1017/epi.2024.54 Published online by Cambridge University Press

should have a *plurality* of epistemic ideals: my revised ideal in some contexts and the Traditional Ideal in others.

My reply

First, to clarify the nature of my revised ideal, this figure *does* aim to acquire positive epistemic states and avoid negative ones – that is, the most valuable element of the Traditional Ideal. This is because having a commitment to epistemic goods crucially involves aiming to realize those epistemic goods: I cannot be committed to epistemic goods if I do not aim to realize them. My discussion has shown, however, that this is not *all* a commitment to epistemic goods involves; it also requires the acknowledgment of error, and this in turn requires accepting the inescapability of error even in ideal conditions. The appealing features of the Traditional Ideal are therefore still captured by the revised ideal – aiming to realize epistemic goods, such as true belief, justification, and understanding – but much more is captured as well. There seems to be no reason, then, for retaining the Traditional Ideal – even if we retain it as just one ideal among others – given that my revised ideal captures what we seem to find most valuable in the former while also revealing further key dimensions of epistemic ideality.[26]

Objection 5

Why not think that the intellectual virtue of open-mindedness captures what is valuable about this practice of changing your mind? Note, for example, the description of Ravitch in my opening examples. Her open-mindedness is specifically praised. By "open-mindedness" here, we can follow Wayne Riggs's (2016) and (2019) accounts, which in part build on Baehr (2011). Riggs argues that one is open-minded insofar as one is willing and able to grasp the interanimating commitments that make up an alternative perspective for making sense of the world. We might think that anyone who changes their minds in the way I am highlighting is simply demonstrating this virtue of open-mindedness.

My reply

First, recall my earlier observations that the practice I am drawing our attention to addresses the *prior* question of an agent's motivation. It gives us crucial evidence relevant to the question of whether our target agent has a commitment to epistemic goods *at all*. Zagzebski implicitly acknowledges this prior question when she writes that "[t]he motivation for knowledge" underlies "the individual intellectual virtues: open-mindedness, fair-mindedness, intellectual flexibility, and so on" (1996: 181).

Second, one can change their mind in the way I am emphasizing and not be open-minded. Suppose, for example, I am a close-minded scientist: I observe other methodologies or paradigms in my field, and I only see misguided and ignorant science.[27] I arrive at this conclusion precisely because I only interact with individual claims from these paradigms abstracted from the broader perspective in which they are embedded and do not grasp how they mutually inform one another, giving the claims an increased epistemic strength collectively that they lack in insolation. I have no interest in doing the work necessary for properly grasping this perspective. But, as it turns out, proponents of this alternative paradigm end up producing results that contradict my own on a crucial set of issues where I have publicly and stridently advocated for a particular view. Once these results are replicated and settled, however, I endorse them and concede that I was wrong. In this case, I change my mind on the basis of the evidence and at a cost to myself, but I do

[26]Thank you to an anonymous reviewer for raising this point. A more detailed account of this pluralism of epistemic ideals may well be worth developing.

[27]See Battaly (2018a and 2018b) for detailed discussion of closed-mindedness.

https://doi.org/10.1017/epi.2024.54 Published online by Cambridge University Press

so without being open-minded. If I were to act in this way and despite my closed-mindedness, then this would still be good evidence of my commitment to epistemic goods.

We might want to say that I *must* have been open-minded to endorse the evidence that I was wrong. But this seems too strong. First, it expands the sense of "open-mindedness" beyond what Riggs or Baehr advocates for. Second, it collapses open-mindedness into responsiveness to the evidence *simpliciter*, and this again seems far too capacious for a specific intellectual virtue. Open-mindedness may well aid someone in being properly responsive to the evidence, but it is not necessary for this responsiveness. As long, then, as one can change one's mind at a likely cost to oneself without evincing open-mindedness, this virtue and the notion I highlight are not equivalent.

Objection 6

It might be argued that I have only shown that error is a regrettable precondition for the excellence of changing one's mind, but not that this precondition is *itself* part of the excellence. Compare: forgiveness and compassion are morally excellent and depend for their existence on agents wronging one another and suffering. But this is not a good argument for an ideal moral life involving wronging and suffering. *Mutatis mutandis* for my arguments about an epistemically ideal figure and the necessity of committing errors.[28]

My reply

I do not accept the claim that an ideal moral life is one without wronging or suffering. This is easier to see in the case of suffering. After all, many of us would choose not to enter Nozick's experience machine, and this is because we think there is something more worthwhile about genuine reality with all its risks of suffering to a simulated, but maximally pleasurable existence. More to the point, though, if forgiveness and compassion cannot be exercised without wronging and suffering, then these phenomena will have to exist, at least to some degree, for there to be ideal moral agents. And it seems clear that moral agents without virtues of forgiveness and compassion are less ideal than moral agents who have them. But even if one wanted to argue that these moral excellences are severable from their preconditions, this is not the case for the specific epistemic practice of changing one's mind. In this case, the error *must* be produced by the agents themselves, or else they will not count as having changed their mind, while in moral cases, the wronging and suffering are presumably caused by factors that have nothing to do with the agents themselves. If one wanted to argue for a clean separation between the preconditions for the excellence and the excellence itself, then, in these moral cases, it might be possible. But there is no way to separate out the epistemic excellence of changing one's mind in light of one's errors from the condition of the agent themselves having committed those errors.

It might nonetheless be argued that I have opened a worrying door here. Consider an analogy with the problem of evil. Some philosophers attempt to solve the latter by arguing that God is justified in allowing the most heinous evils – genocide, torture of children, etc. – because they also enable the instantiation of exemplary moral conduct, such as heroically saving individuals from such conditions, which in turn generate a better moral world than one without these evils. But many philosophers do not find this line compelling, even viewing it as an appalling rationalization of atrocity. (I count myself among them.) But, by the same token, if we do not accept that great evil is justified because it facilitates the existence of moral goods, then we should not accept that great epistemic failures are justified or fixtures of an ideal epistemic world because they facilitate the existence of a commitment to epistemic goods.

[28]I borrow this formulation of the objection from an anonymous reviewer's helpful comments.

https://doi.org/10.1017/epi.2024.54 Published online by Cambridge University Press

A first response might be to deny the relevance of the analogy: the implausibility of this theodicy derives from the distinctly moral character of these wrongs, but epistemic failures do not have this moral weight and therefore can be compensated for enabling a world with commitments to epistemic goods. But I prefer a different reply that harkens back to my reply to Objection 2 and the notion of the epistemic masochist: allowance for some epistemic failure does not license all forms of epistemic failure. That is, a world with a degree of epistemic failure does not entail rampant epistemic irrationality. In fact, it is hard to see how such a world would be compatible with agents who are committed to the existence of epistemic goods: such agents would be perpetually and arbitrarily mistaken, seemingly lacking any commitment to epistemic goods.

But a certain degree of epistemic failure is compatible with an ideal epistemic world. By analogy, limited forms of suffering or wronging seem compatible with an ideal moral world, though in a way that importantly rules out the above theodicy (a happy result in my view). The pain of heartbreak or the wrong of breaking a promise to a friend opens up opportunities for the exercise of compassion and forgiveness; these are forms of suffering and wronging that therefore plausibly generate a more ideal moral world by facilitating these virtues. Epistemic failures not born of unchecked irrationality similarly generate the conditions for a more ideal epistemic world overall – one in which agents can be committed to epistemic goods.

Objection 7

We might question why the practice of changing one's mind deserves praise at all. The very same figures I cite might be criticized for being fickle or "flip-floppers." In fact, Putnam himself notes frequently encountering this charge.

My reply

As phrased, this objection denies that changing one's mind in response to the evidence and at a likely cost to oneself is *ever* praiseworthy. This seems clearly wrong. Even our objector will admit that some cases of changing one's mind are not mere instances of fickleness or flip-flopping. They may dislike my examples, but then they are free to insert their own.

But the objector might moderate their position. They might allow that there are genuinely praiseworthy cases of changing one's mind in this way, but the vast majority of cases are not praiseworthy (including perhaps some or all the ones I cite). It could be argued that most cases require what Elise Woodard has recently called "epistemic atonement" or "making up for one's previous epistemic failures, including believing badly," rather than any type of praise (2023: 163). To epistemically atone, Woodard argues that agents "must restore *epistemic* trust and demonstrate *epistemic* trustworthiness, or trustworthiness with respect to their doxastic states and processes," which they can do through a variety of practices, such as "accepting responsibility, offering explanations, expressing negative attitudes, offering forward-looking commitments, and engaging in community service" (172 and 168).

But as long as this objection concedes that *some* cases of changing one's mind in the way I highlight are praiseworthy, then my account stands.[29] Furthermore, I am skeptical of elements of Woodard's account of epistemic atonement. If I correct my epistemic mistakes, then it seems that I am doing just what I ought to do as a learner and epistemic

[29]Elsewhere, Woodard agrees with this assessment: "[W]e tend to regard agents who carefully redeliberate and change their opinions as particularly thoughtful, open-minded, and responsible epistemic agents," citing Putnam as an example (2022: 328).

https://doi.org/10.1017/epi.2024.54 Published online by Cambridge University Press

agent. It is odd to demand that students (which we all are for at least some domains) must perpetually atone for these errors.

Woodard's two main examples are revealing. The first is Biden's reversal of his stance on the Hyde Amendment, a move widely criticized by many Democrats who nonetheless agree with Biden's reversed stance. For Woodard, "[t]he central problem, on my diagnosis, is that Biden failed to acknowledge or make up for his previous epistemic failures" (165). I see these cases differently. The criticisms Woodard cites seem better read as expressions of frustration at politicians' engaging in shamelessly self-interested, epistemically indifferent behavior. That is, the frustration with Biden is not that he fails to explain his epistemic failures but that he does not seem to be in the epistemic domain *at all*: he is simply opting for a now popular view because it serves his interests to do so, and the cited criticisms seem more plausibly interpreted as (justified) irritation at the pervasiveness of Frankfurtian bullshit in politics.[30]

Woodard's second example concerns Marjorie, who was once a 9/11 "Truther" despite exposure to the evidence that her beliefs were false and who only abandoned these beliefs recently: "Are your concerns about her epistemic inclinations fully assuaged by the fact that she has since changed her mind? For many, the answer to these questions is *no*. When you first learn about her doxastic track record, you would likely regard her with some epistemic suspicion" (167). I do not share the intuition that we would view Marjorie with epistemic suspicion. If anything, I would be impressed with Marjorie's epistemic character. Her ability to escape a highly epistemically polluted, echo-chambered environment and arrive at better views is an extremely difficult task that plausibly requires *superior* epistemic capacities. Marjorie has not just contingently ended up in a healthy epistemic environment and with the "right" views, as the non-echo-chambered have. She must have exceptional epistemic abilities to both see through and then exit this noxious environment. I am not alone in this judgment. We often praise individuals who escape echo chambers, and we view them as not just moral, but epistemic exemplars.[31]

As I read them, Woodard's examples reveal that we are suspicious of changes in mind when they are advantageous for the agent (as in Biden's case) or when they are not obviously in response to the evidence (as in versions of Marjorie's case). Our hesitance to regard these changes as genuine, in turn, illuminates the weightiness in our epistemic lives of the practice of changing one's mind in response to the evidence. We do not want to concede that certain cases are genuine, rather than being driven by self-interest or fickleness, because if they *were* genuine, then we would have to accept that we are engaging with an agent who has a commitment to epistemic goods, and this is not a concession we are always prepared to make lightly. When we treat their activity as the product of a commitment to epistemic goods, we no longer view the agent as, for example, an adversary in a political competition whose apparent reasons, arguments, and evidence are beside the point; we now view them as a fellow inhabitant in the space of reasons to whom we are accountable.[32]

[30]Woodard seems to agree with this reading at points, claiming that what agents in these circumstances will need to do is "affirm their status as *epistemic* agents, indicating that they are responsive to epistemic reasons and criticism" (175).

[31]See, for example, Nguyen's paradigm case of echo chamber escape, Derek Black (Nguyen 2020).

[32]Thank you to audiences at Beijing Normal University at Zhuhai (and to M Dentith for the invitation) and Wake Forest University for feedback on this paper. Many thanks to Andrew Howat, Brian Klug, Francisco Gallegos, Johnny Dixon, Tobias Flattery, Christian Miller, and Devin Lane for very helpful comments on drafts. I am grateful to Hailey Huget for discussing these ideas with me at length as well as reading and commenting on multiple drafts. A special thanks to an anonymous reviewer at *Episteme*, who gave some of the most helpful comments I have ever received, prompting me to rethink the paper in crucial places and substantially improving the final product.

https://doi.org/10.1017/epi.2024.54 Published online by Cambridge University Press

References

Annas J. (2011). *Intelligent Virtue*. Oxford: Oxford University Press.
Baehr J. (2011). *The Inquiring Mind: On Intellectual Virtues and Virtue Epistemology*. Oxford: Oxford University Press.
Bardon A. (2020). *The Truth About Denial: Bias and Self-Deception in Science, Politics, and Religion*. Oxford: Oxford University Press.
Battaly H. (2018a). 'Can Closed-Mindedness be an Intellectual Virtue?' *Royal Institute of Philosophy Supplement* **84**, 23–45.
Battaly H. (2018b). 'Closed-mindedness and Dogmatism.' *Episteme* **15**(3), 261–82.
Calhoun C. (2009). 'What Good Is Commitment?' *Ethics* **119**(4), 613–41.
Carter J.A., & McKenna R. (2020). 'Skepticism Motivated: On the Skeptical Import of Motivated Reasoning.' *Canadian Journal of Philosophy* **50**(6), 702–18.
Chang R. (2013). 'Commitments, Reasons, and the Will.' In R. Shafer-Landau (ed.), *Oxford Studies in Metaethics, Volume 8*, pp. 74–113. Oxford: Oxford University Press.
DiPaolo J. (Forthcoming). "I'm, Like, a Very Smart Person': On Self-Licensing and Perils of Reflection.' *Oxford Studies in Epistemology*.
DiPaolo J. (2021). 'The Word of a Reluctant Convert.' *Synthese* **198**(1), 557–82.
Greco J. (2021). "Pritchard's Case for Veritism." *Epistemology & Philosophy of Science* **58**(4), 46–53.
Guerrero A.A. (2017). 'Living with Ignorance in a World of Experts.' In R. Peels (ed.), *Perspectives on Ignorance From Moral and Social Philosophy*, pp. 156–185. New York: Routledge.
Herman B. (1981). 'On the Value of Acting from the Motive of Duty.' *Philosophical Review* **90**(3), 359–82.
Hornsey M.J., & Fielding, K.S. (2017). 'Attitude Roots and Jiu Jitsu Persuasion: Understanding.' *American Psychologist* **72**(3), 459–73.
Kamtekar R. (2004). 'Situationism and Virtue Ethics on the Content of Our Character.' *Ethics* **114**(3), 458–91.
Kant I. (1786/2011). *Groundwork of the Metaphysics of Morals*, ed. and trans. M. Gregor and J. Timmerman. Cambridge: Cambridge University Press.
Kornblith H. (1999). 'Distrusting Reason.' *Midwest Studies In Philosophy* **23**(1), 181–96.
Korcz K.A. (2021). 'The Epistemic Basing Relation.' *The Stanford Encyclopedia of Philosophy*. https://plato.stanford.edu/archives/spr2021/entries/basing-epistemic/
Kvanvig J.L. (2005). 'Truth is Not the Primary Epistemic Goal. In M. Steup and E. Sosa (eds.), *Contemporary Debates in Epistemology*, pp. 285–295. Malden: Blackwell.
Levy G. (2020). 'Peter Beinart's Great Change.' *Haaretz*. https://www.haaretz.com/opinion/2020-07-12/ty-article-opinion/.premium/peter-beinarts-great-change/0000017f-e987-da9b-a1ff-edef289c0000
McBrayer J.P. (2024). '*Caveat Auditor*: Epistemic Trust and Conflicts of Interest.' *Social Epistemology* **38**(3), 290–301.
McDowell J. (1979). 'Virtue and Reason.' *The Monist* **62**(3), 331–50.
McKenna R. (2023). *Non-Ideal Epistemology*. Oxford: Oxford University Press.
Nguyen C.T. (2020). 'Echo Chambers and Epistemic Bubbles.' *Episteme* **17**(2), 141–61.
Nietzsche F. (1886/1989). *Beyond Good and Evil*, trans. Walter Kaufmann. New York: Vintage Books.
Nussbaum M.C. (2016). 'Hilary Putnam (1926–2016).' *Huffington Post*. https://www.huffpost.com/entry/hilary-putnam-1926-2016_b_9457774
Putnam H. (1988). *Representation and Reality*. Cambridge: The MIT Press.
Roberts R.C., & Wood W.J. (2007). *Intellectual Virtues: An Essay in Regulative Epistemology*. Oxford: Oxford University Press.
Rovelli C. (2020). *There are Places in the World Where Rules Are Less Important Than Kindness: And Other Thoughts on Physics, Philosophy, and the World*. London: Penguin Random House Ltd.
Riggs W. (2016). 'Open-Mindedness, Insightfulness, and Understanding.' In J. Baehr (ed.), *Intellectual Virtues and Education: Essays in Applied Epistemology*, pp. 18–37. New York: Routledge.
Riggs W. (2019). 'Open-Mindedness.' In H. Battaly (ed.), *The Routledge Handbook of Virtue Epistemology*, pp. 141–154. New York: Routledge.
Schulz K. (2010). *Being Wrong: Adventures in the Margin of Error*. New York: HarperCollins Publishers.
Shaver R. (2023). 'Egoism.' *Stanford Encyclopedia of Philosophy*. https://plato.stanford.edu/archives/spr2023/entries/egoism/
Shields M. (2022). 'Rethinking Conspiracy Theories.' *Synthese* **200**(4), 1–29.

https://doi.org/10.1017/epi.2024.54 Published online by Cambridge University Press

Tenenbaum S. (2021). *Rational Powers in Action: Instrumental Rationality and Extended Agency*. Oxford: Oxford University Press.

Thompson C. (2018). 'Real Heroes Have the Guts to Admit They're Wrong.' *Wired*. https://www.wired.com/story/real-heroes-have-the-guts-to-admit-theyre-wrong/

Toole B. (2022). 'Demarginalizing Standpoint Epistemology.' *Episteme* **19**(1), 47–65.

Tosi J., & Warmke B. (2022). 'Don't Block the Exits.' In J.P. Messina (ed.), *New Directions in the Ethics and Politics of Speech*, pp. 50–60. New York: Routledge.

Wolfe A. (2010). 'The Education of Diane Ravitch.' *The New York Times*. https://www.nytimes.com/2010/05/16/books/review/Wolfe-t.html

Woodard E. (2023). 'Epistemic Atonement.' In R. Shafer-Landau (ed.), Oxford Studies in Metaethics Volume 18, pp. 163–190. Oxford: Oxford University Press.

Woodard E. (2022). 'A Puzzle about Fickleness.' *Noûs* **56**(2), 323–42.

Worsnip A. (2023). 'Suspiciously Convenient Beliefs and the Pathologies of (Epistemological) Ideal Theory.' *Midwest Studies in Philosophy* **47**, 237–68.

Zagzebski L. (1996). *Virtues of the Mind: An Inquiry into the Nature of Virtue and the Ethical Foundations of Knowledge*. Cambridge: Cambridge University Press.

Matthew Shields is an assistant professor of philosophy at Wake Forest University. He specializes in epistemology, philosophy of language, and metaphilosophy.

Cite this article: Shields M. (2026). "On the value of changing your mind." *Episteme* **23**, 388–410. https://doi.org/10.1017/epi.2024.54

https://doi.org/10.1017/epi.2024.54 Published online by Cambridge University Press