Episteme (2026), 1–8
doi:10.1017/epi.2026.10161

episteme

ARTICLE

Mind-Changing and Epistemic Commitment without Error: A Reply to Shields

Oushinar Nath

University College London, UK
Email: oushinar.nath.17@ucl.ac.uk

(Received 4 August 2026 UTC; revised 2 September 2026 UTC; accepted 29 September 2026 UTC)

Abstract

Traditional conceptions of the epistemically ideal agent emphasise the possession of positive epistemic states (e.g., truth, knowledge, understanding, etc.) and the avoidance of falsehood and other epistemic defects. Matthew Shields challenges this picture and argues that flawless epistemic performance may be merely self-serving: an agent might pursue epistemic goods only because achieving them produces personal satisfactions. By contrast, changing one's mind in response to good evidence and at a likely cost deprives the agent of such satisfaction. It, therefore, provides especially strong evidence that the agent values epistemic goods for their own sake. From this, Shields concludes that an epistemically ideal agent must have had at least one false belief or defective practice which they subsequently corrected. I argue that although Shields correctly identifies an inadequacy in accounts of ideal epistemic agency that focus exclusively on epistemic outcomes, it does not follow that the ideal agent must actually have been wrong – what matters is not so much the correction of error, but the agent's practical willingness to expose their epistemic position to possible correction when doing so is costly.

Keywords: Mind-changing; epistemic commitment; epistemic selflessness; epistemic ideality; open-mindedness

1. Shields on mind-changing and epistemic commitment

Shields's argument begins from what he calls the 'Problem of Mere Epistemic Performance'. An agent may display impeccable epistemic conduct without possessing a genuine commitment to epistemic goods. She may consistently form true beliefs, reason competently, respond appropriately to evidence, and exhibit the intellectual virtues while doing so only because epistemic success happens to serve some non-epistemic interests. Shields illustrates this possibility through the *Self-Serving Know-It-All*: an agent whose beliefs are always true and appropriately justified, but whose epistemic impeccability is sustained by the satisfaction she derives from being epistemically impeccable. Were epistemically defective conduct instead to serve her interests better, she would abandon her present practices. Her relationship to epistemic goods is

© The Author(s), 2026. 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.2026.10161 Published online by Cambridge University Press

therefore merely contingent (Shields 2026: 395–97). For Shields, epistemic ideality accordingly requires more than mere flawless epistemic performance – it requires a genuine, non-contingent commitment to epistemic goods. Following Calhoun (2009), Shields understands commitment as something that must be manifested and sustained through action rather than merely postulated as a psychological state. Such commitment becomes particularly revealing when acting in accordance with epistemic goods conflicts with the agent's non-epistemic interests. This explains the special significance Shields attributes to costly self-correction. Admitting that one was mistaken may involve professional, reputational, material, social, or psychological costs, including the loss of the satisfaction associated with being right. An agent who nevertheless corrects her mistake in response to good evidence thereby provides particularly strong evidence that epistemic goods, rather than her non-epistemic interests, govern her conduct. Shields describes this broader phenomenon as 'epistemic selflessness': the prioritisation of epistemic goods over one's own non-epistemic interests (2026: 398). Shields's argument can, therefore, be constructed as follows:

(P1) The epistemically ideal agent must possess a non-contingent commitment to epistemic goods.
(P2) Such a commitment must be manifested in conduct when epistemic and non-epistemic interests conflict.
(P3) For an otherwise epistemically flawless agent, the only conduct capable of adequately manifesting such commitment is costly self-correction.
(P4) Costly self-correction requires a prior false belief or deficient epistemic practice.
(C) Therefore, the epistemically ideal agent must have committed and subsequently corrected at least one epistemic error.

My objection here concerns (P3). Shields's conclusion requires not merely that costly self-correction is an especially powerful way of manifesting epistemic commitment, but that it is the *only* adequate way available to an otherwise epistemically flawless agent.[1] I shall argue that this exclusivity claim is false – an agent may also incur substantial costs in deliberately exposing an otherwise well-supported epistemic attitude to possible correction, even when no error is ultimately discovered and no change of mind is required.

2. Correction, exposure, and costs

It will be helpful here to distinguish two ways in which an agent's commitment to epistemic goods can be manifested. I shall call the first, *costly self-correction*: an agent

[1]Shields sometimes formulates his central claim comparatively – changing one's mind in response to good evidence and at a likely cost provides the *best* evidence of an agent's commitment to epistemic goods (2026: 395, 399–400). However, he forcefully concludes that the epistemically ideal agent must have committed and corrected an error, since such failures are *necessary* to establish that commitment (2026: 399–401). For this necessity conclusion to follow, however, changing one's mind under these conditions must not merely be the best evidence of commitment but the only epistemically adequate means by which an otherwise flawless agent can manifest it. If an error-free form of conduct can adequately manifest the same commitment, that necessity conclusion does not follow, even if changing one's mind remains the strongest evidence of commitment.

https://doi.org/10.1017/epi.2026.10161 Published online by Cambridge University Press

engages in costly epistemic self-correction when she discovers that some existing belief or epistemic practice is defective and revises it despite significant incentives to retain it. These are the cases on which Shields primarily focuses. But consider a second way, which I shall call *costly self-exposure*: An agent engages in costly epistemic self-exposure when she deliberately subjects an existing epistemic attitude to serious potential defeat and bears significant costs in doing so, even though the resulting inquiry may ultimately discover no error requiring correction. Here, the agent does not merely claim that she would change her mind were counterevidence somehow to appear. Rather, she actively takes steps that make the discovery of such evidence possible and, sometimes, more likely. Such kind of conduct resembles what psychologists describe as 'active open-minded thinking'. On this conception, open-mindedness does not consist merely in a passive willingness to change one's mind should contrary evidence arise. Rather, open-minded agents actively seek out alternative possibilities, consider reasons against their existing views, and engage seriously with those alternatives with a willingness to revise their attitudes when revision is warranted (e.g., Baron 2019; Stanovich and Toplak 2023). Related ideas appear in philosophical literature too. Medina (2013: 33–34), for example, criticises agents who wilfully insulate themselves from perspectives that conflict with their own, whether through intellectual laziness or because 'one doesn't need to know'. Similarly, Peters *et al.* (2025: 3–4) argue that agents embedded in echo chambers fail to be appropriately open-minded insofar as they are not positioned to encounter and engage seriously with opposing views. What these discussions suggest is that epistemic responsibility may require more than just responsiveness to evidence one already possesses – it may require agents to place themselves in circumstances in which potentially challenging evidence can become available. Costly self-exposure is a particularly demanding instance of this more active form of open-mindedness: the agent does not merely leave open the possibility of future belief revision should countervailing evidence arise, but instead, actively seeks criticism, examines possible defeaters, consults dissenting experts, structures her inquiry so that weaknesses in her epistemic position are more likely to be revealed, and so on – she actively arranges her epistemic environment in ways such that relevant counterevidence becomes more likely to be available. When doing so carries substantial professional, material, social, or psychological costs, this exposure can manifest precisely the sort of non-contingent commitment to epistemic goods that Shields identifies.

Now, this distinction between self-correction and self-exposure matters because the conflict between epistemic goods and non-epistemic interests need not arise only after an error has already been identified, as Shields's argument requires. It can arise earlier, at the stage at which an agent must decide whether to expose a presently well-supported and personally advantageous belief to further scrutiny. Doing this may require her to sacrifice many of her non-epistemic interests, like time, resources, priority, reputation, psychological comfort etc., in order to inquire and make potential errors discoverable. In such circumstances, closing the inquiry may serve those interests much better. Nevertheless, epistemic responsibility may require her to further investigate those matters. Costly self-exposure, therefore, resembles Shields's cases in the respect that matters most for his conception of epistemic commitment – it places an epistemic demand in conflict with the agent's non-epistemic interests. The difference concerns the stage at which the conflict occurs. In costly self-correction, the agent has already encountered evidence requiring revision and must decide whether to respond to it appropriately. In costly self-exposure, on the other hand, the agent must decide whether to place herself in a position in which such evidence could be discovered in the first

https://doi.org/10.1017/epi.2026.10161 Published online by Cambridge University Press

place. Importantly, unlike costly self-correction, in costly self-exposure such costs are incurred regardless of whether further inquiry eventually reveals an error. In fact, further investigation may instead vindicate the agent's original position. Such conduct can satisfy the following condition on epistemic commitment:

> (*Costly Commitment Condition*): When epistemic requirements conflict with an agent *S*'s non-epistemic interests, *S*'s pursuit of epistemic goods continues to govern *S*'s conduct.

Costly self-correction is one way of satisfying this condition, but costly self-exposure may be another. Shields' own category of epistemic selflessness supports this possibility. He characterises epistemic selflessness as consisting in prioritising epistemic goods over one's non-epistemic interests, and acknowledges that it encompasses a broader class of conduct than costly self-correction alone (2026: 398). This then raises the question: why must the conflict between epistemic goods and non-epistemic interests occur only after an error has been identified, rather than when an agent decides, say, whether to place a highly valued belief at risk of defeat?

3. The corrigible know-it-all

Consider the following case:

> *(Historian)*: Helen is a historian who has spent several years researching the collapse of a twentieth-century government. On the basis of recently declassified documents, private correspondence, and contemporaneous diaries, she concludes that a senior official who was long thought to have supported the regime until its final days had in fact secretly begun cooperating with the opposition several months earlier. She sends the relevant archival materials and a draft of her book containing her argument to several historians for independent assessment. This includes Julio, who happens to be the most prominent historian working on that period. Whereas all the other historians replied that Helen's sources are authentic and authoritative, and her interpretation is well supported by them, Julio does not reply. However, in a recent public lecture, when he is asked about Helen's forthcoming book, Julio remarks, 'I've looked through some of the material, and I'm not convinced that the official began cooperating with the opposition as early as Helen claims.' He, however, does not explain why he remains unconvinced.
>
> One of Helen's colleagues later mentions Julio's remarks to her and asks what she thinks. Helen has known Julio for many years and regards him as highly knowledgeable and reliable in his historical judgments. She also knows, however, that Julio is working simultaneously on several projects and that he sometimes comments on archival material before he has examined all of it in detail. Helen replies, 'Look, the documents have been independently verified, and the other historians I sent the material to agree that they support my interpretation. So, I do believe that the official was already cooperating with the opposition at an early stage. Julio is, of course, the most respected historian working on this period, and what he says needs to be taken very seriously. But he hasn't yet replied to my emails about the documents, and he hasn't explained in the lecture why he thinks my

https://doi.org/10.1017/epi.2026.10161 Published online by Cambridge University Press

> interpretation is wrong. Still, if he can provide any convincing counterevidence, I'm open to revising the argument, and even changing my mind.'
>
> Investigating Julio's concerns, however, would come at a significant personal and professional cost to Helen. Her book is already scheduled for publication, and delaying it may allow another historian to publish related findings first. It may also weaken the publicity surrounding the book and jeopardise her chances of receiving professional recognition for the discovery. Nevertheless, Helen asks the publisher to postpone publication, contacts Julio again to ask what his reservations might be based on, and arranges to examine the archival collection again. Julio is in fact wrong – he had overlooked the dating of one of the documents and had not examined some pieces of corroborating evidence – and Helen is correct.

Here, Helen has formed the belief that the official was already cooperating with the opposition from an earlier stage by reviewing the available first-order historical evidence in a reliable manner. For instance, Helen herself is an expert on the relevant historical period. Given her expertise, she might have examined recently declassified government records, private correspondence, contemporaneous diaries, and other archival materials with considerable care, cross-checking dates, provenance, and contextual details across multiple independent sources. She may also have assessed alternative interpretations of the documents and compared them with the established historical record in a systematic manner using her expert historian's skills. Helen had additionally sent the relevant archival material and a draft of her argument to other historians working on the same area, who independently examined the documents and checked whether her interpretation was supported by the available evidence. Since they confirm her conclusion, this provides Helen with very strong higher-order evidence in support of her epistemic position. Accordingly, (assuming fallibilism), it is safe to say that Helen *knows*, more specifically has first-order knowledge, that the official was already cooperating with the opposition at an earlier stage.

But even though Julio's testimony does not defeat Helen's first-order knowledge that the official was already cooperating with the opposition at that stage, it still is epistemically significant for her. Given that Julio is the most prominent expert in the area and is highly reliable, his disagreement provides Helen with sufficiently strong countervailing higher-order evidence. This in turn produces certain epistemic reasons for her to be disposed to change her mind.[2] More specifically, it could be said that, given her position as an expert in history, she has developed a reliable sensitivity towards what she ought to consider as evidentially relevant bearing on historical matters. She might think that Julio may be in possession of certain archival materials that she does not have access to, that he may have noticed details that she has overlooked, that he may have considered other hypotheses that she has not, etc. Now, Helen need not be aware of what precisely such missing possibilities are or could be. However, given her evidential sensitivity qua expert, she takes these unaware possibilities as being evidentially relevant for figuring out whether her historical judgement about the official is correct. Owing to this, she is in an epistemic position where even though she has first-order knowledge that the official was conspiring at an early stage, she is in an impoverished higher-order

[2]It needs to be borne in mind that what I am claiming here is that the countervailing evidence is sufficiently strong for Helen to remain *open* to changing her mind, without being sufficiently strong for her to *change* her mind.

https://doi.org/10.1017/epi.2026.10161 Published online by Cambridge University Press

epistemic state regarding that piece of knowledge, such that continued inquiry into the matter is still epistemically appropriate. It is this impoverished higher-order epistemic position that also disposes her to change her mind.

Now, by stipulation, Helen is epistemically responsible. But an epistemically responsible agent does not merely respond appropriately to the evidence she presently possesses; she must also remain appropriately sensitive to indications that her evidential position may be impoverished. Since she recognises that Julio may possess relevant evidence, or may have considered possibilities that she has not, by her own lights, she cannot responsibly regard the inquiry into the matter as completely settled.[3] This means that she treats her belief as potentially revisable in light of further inquiry. The relevant disposition is, therefore, conditional rather than immediately doxastic – Helen need not now change her mind, but rather, she ought to be disposed to change her mind conditional on further inquiry revealing that Julio in fact possesses countervailing evidence, has identified a defect in her reasoning, etc. Her impoverished higher-order epistemic position thus prevents her from treating her present epistemic success as sufficient for closing inquiry. To refuse to remain open to changing her mind in these circumstances would amount to being epistemically irresponsible. Thus, the higher-order evidence generated by Julio's testimony rationally supports her disposition towards changing her mind without actually having to change her mind.

Crucially, this disposition for Helen manifests through costly self-exposure. Rather than simply declaring or thinking that she would change her mind if presented with compelling counterevidence, as seen in *Historian*, she takes active and costly steps to determine whether such counterevidence exists – she postpones publication, reopens the relevant parts of her inquiry, seeks clarification from Julio, and subjects her interpretation to further scrutiny despite the professional and personal costs involved. Her disposition to change her mind is therefore practically manifested in her willingness to make her present belief vulnerable to correction. Costly self-exposure, in this sense, is her practical expression of the epistemic reasons generated by the higher-order evidence from Julio's testimony.

Helen's conduct here is significant because the relevant conflict between epistemic goods and non-epistemic interests has already occurred, even though she has not discovered, and may ultimately never discover, an error. She could have protected the benefits associated with her belief by dismissing Julio's unsupported reservations and proceeding with publication. Instead, she deliberately makes her epistemic position vulnerable to correction. In doing so, she allows epistemic considerations to govern her conduct when they conflict with her professional and personal interests. Helen, therefore, satisfies the *Costly Commitment Condition*: when epistemic requirements conflicted with her non-epistemic interests, her pursuit of epistemic goods continued to govern what she did. Her conduct also exemplifies Shields's conception of epistemic

[3]One might wonder whether the cognitive profile described here is better understood in terms of intellectual humility rather than a disposition to change one's mind. The two may overlap insofar as both can involve appropriate sensitivity to one's cognitive limitations. Nevertheless, intellectual humility by itself does not entail a willingness to revise one's beliefs. Peters *et al.* (2025: 3), for example, argue that an intellectually humble agent may recognise her limitations without thereby being disposed either to engage seriously with opposing considerations or to alter her existing beliefs. The cognitive profile relevant here involves something further: Helen does not merely acknowledge her epistemic limitations – she takes Julio's testimony to be a potential source of counterevidence and remains disposed to change her historical belief should further inquiry reveal sufficient reason for doing so. What matters here, therefore, is not intellectual humility alone, but a disposition towards belief revision in response to epistemically relevant considerations.

https://doi.org/10.1017/epi.2026.10161 Published online by Cambridge University Press

selflessness: she prioritised responsible inquiry over the benefits she would have reaped from treating her historical conclusion as settled. Moreover, this commitment is manifested through action. Unlike Shields's paradigmatic agents, however, Helen is not required to change her mind, for further inquiry may instead vindicate her original conclusion. *Historian*, therefore, demonstrates that costly self-correction is not the only way to manifest a non-contingent commitment to epistemic goods; costly self-exposure to possible correction can do so as well. If this is correct, then actual error is not necessary for the kind of epistemic commitment that Shields takes to be required by ideal epistemic agency.

At this juncture, Shields might object that Helen still gets to retain the satisfaction of being right; she might pursue scrutiny because she enjoys viewing herself as exceptionally rigorous or open-minded. Costly self-exposure, therefore, does not eliminate self-serving explanations as effectively as admitting that one was wrong. In response, it could be said that costly self-exposure need not eliminate every possible self-serving motivation in order to undermine Shields's conclusion. Costly self-correction does not do so either – after all, an agent might derive satisfaction from correcting her own mistake. What matters instead is that Helen incurs substantial costs and deliberately increases the likelihood that a personally advantageous belief will be undermined. Her conduct, therefore, supplies practically manifested evidence of epistemic commitment, even if costly self-correction may sometimes supply stronger evidence. The existence of an epistemically adequate alternative is sufficient to block the claim that epistemic ideality requires prior error.[4] Helen, thus, exemplifies a more general epistemic figure:

> (*Corrigible Know-It-All*): They respond to every epistemic phenomenon as they ought to. Their beliefs are always true and always have the best justifications for them. They possess all the intellectual virtues, and they assess their evidence as they ought to. However, although they know the relevant propositions, they do not take themselves to possess all the evidence that may bear on them. They recognise that there may be evidentially relevant possibilities of which they are unaware. Since they do not know what such unpossessed evidence would be, or whether it would ultimately confirm or undermine their existing beliefs, its mere possibility does not defeat their first order knowledge. Nevertheless, they remain reliably sensitive to indications that further inquiry may disclose evidence requiring them to change their minds. They therefore seek relevant criticism, take competent disagreement seriously, and remain disposed to change their minds should further evidence warrant doing so. Where such inquiry conflicts with their non-epistemic interests, they are willing to incur the relevant costs of making their epistemic position vulnerable to correction. As it happens, further inquiry always vindicates their existing beliefs, and they therefore never actually need to change their minds.

Both my *Corrigible Know-It-All* and Shields's *Self-Serving Know-It-All* may satisfy the traditional first-order requirements of epistemic ideality, but only the former has her commitment to epistemic goods practically tested when epistemic goods conflict with non-epistemic self-interests. Importantly, this does not require the correction of an

[4]This should not be understood as a rejection of (P4) in Shields's argument. (P4) maintains that prior error is necessary for *self-correction* (I am, in fact, happy to grant that claim). My objection concerns a stronger thesis, viz., that prior error is necessary for adequately manifesting a non-contingent *commitment to epistemic goods*.

https://doi.org/10.1017/epi.2026.10161 Published online by Cambridge University Press

actual error. The *Corrigible Know-It-All* can demonstrate her willingness to discover and respond to possible defects in her epistemic position by making that position vulnerable to further scrutiny. This is not to say that costly self-correction lacks epistemic significance. Shields may be right that correcting oneself at substantial cost provides particularly strong evidence of epistemic commitment. What does not follow from this, however, is that the ideal epistemic agent must previously have been mistaken. Costly self-correction and costly self-exposure are both ways of manifesting commitment under conditions of conflict, but only the former requires prior error. Epistemic ideality may, therefore, require corrigibility under conditions of cost, rather than a history of error and subsequent correction.

References

Baron J (2019). 'Actively Open-Minded Thinking in Politics.' *Cognition* **188**, 8–18.

Calhoun C (2009). 'What Good is Commitment?' *Ethics* **119**(4), 613–41.

Medina J (2013). *The Epistemology of Resistance: Gender and Racial Oppression, Epistemic Injustice, and the Social Imagination.* New York: Oxford University Press.

Peters K, Turner C and Battaly H (2025). 'Intellectual Humility without Open-Mindedness: How to Respond to Extremist Views.' *Episteme* **22**, 1–25.

Shields M (2026) On the Value of Changing Your Mind, *Episteme*, **23**, 388–410.

Stanovich K.E and Toplak M.E (2023). 'Actively Open-Minded Thinking and Its Measurement.' *Journal of Intelligence* **11**(2), 27.

Oushinar Nath is a teaching assistant in Philosophy at University College London. He specialises in epistemology and philosophy of mind, and his interests include wisdom, open-mindedness, inquiring attitudes, mind-changing, and cognitive skills.

Cite this article: Nath O. (2026). "Mind-Changing and Epistemic Commitment without Error: A Reply to Shields." *Episteme* 1–8. https://doi.org/10.1017/epi.2026.10161

https://doi.org/10.1017/epi.2026.10161 Published online by Cambridge University Press