

## CAN EPISTEMIC PATERNALISTIC PRACTICE MAKE US BETTER EPISTEMIC AGENTS?

Giada Fratantonio

School of Humanities  
University of Glasgow

**ABSTRACT.** Can epistemic paternalistic practices make us better epistemic agents? While a satisfying answer to this question will ultimately rest at least partly on empirical findings, considering the epistemological discussion on evidence, knowledge, and epistemic virtues can be insightful. In this paper, Giada Fratantonio argues that we have theoretical reasons to believe that strong epistemic paternalistic practices may be effective at mitigating some evidential mistakes, in fostering true belief, and even for allowing the subject of the intervention to gain knowledge. However, we have reasons to expect that these practices will not be able to make the subject of the intervention an overall better epistemic agent at the dispositional level. She then considers weak epistemic paternalistic practices, e.g., epistemic nudging, and provides some reasons for optimism. Finally, Fratantonio considers the implications that these theoretical considerations have for education.

**KEY WORDS.** epistemic paternalism; epistemic nudging; evidentialism; epistemic vices

## INTRODUCTION

According to a popular view among epistemologists, epistemic justification requires that one believes on the basis of one's supporting evidence.<sup>1</sup> Call this view *evidentialism*.<sup>2</sup> Evidentialism is traditionally motivated by appealing to what is sometimes known as "veritism," the idea that maximizing true beliefs and avoiding false beliefs is what matters from a purely epistemic point of view.<sup>3</sup> Given that believing what one's evidence supports seems to be the best way to achieve this goal, if veritism is true, then it seems we have a good case in favor of evidentialism.

However, although following one's evidence seems to be the best way to get to the truth, it is not an *infallible* way to get to the truth. Sometimes we can believe on the basis of our evidence that supports that *p*, e.g., by making

1. See Earl Conee and Richard Feldman, *Evidentialism* (Oxford: Oxford University Press, 2004). Providing a full evidentialist account of knowledge and justification is a non-trivial task. There's no consensus on how to define what evidence is, what it means for someone to have evidence, what it means for evidence to support a proposition, what it means for someone to believe on the basis of one's evidence. Here, I remain neutral on these issues, although I will assume the popular (though not unproblematic) threshold view of evidential support on which evidence *e* supports a proposition *p* when it makes *p* sufficiently likely.

2. Although very popular, evidentialism is not unanimously endorsed by epistemologists. The traditional competing view is process reliabilism [see Alvin I. Goldman, *Epistemology and Cognition* (Cambridge, MA: Harvard University Press, 1986)].

3. See, for example, Richard Feldman, "The Ethics of Belief," *Philosophy and Phenomenological Research* 60, no. 3 (2000): 667–695; and Alvin I. Goldman and Erik J. Olsson, "Reliabilism and the Value of Knowledge," in *Epistemic Value*, ed. Adrian Haddock, Alan Millar, and Duncan Pritchard (Oxford: Oxford University Press, 2009), 19–41.

$p$  very probable, and yet  $p$  turns out to be false. In other words, unbeknownst to us, sometimes our evidence can be misleading. Other times, it is difficult to appropriately believe on the basis of one's supporting evidence in the first place. There are different ways of committing what I call "evidential mistakes." We can be mistaken about the evidence we have, or what our evidence supports. Sometimes we believe that  $p$  on the basis of evidence that offers only very little support to  $p$ . In other cases, instead, we believe that  $p$ , and it happens that our evidence  $e$  does indeed sufficiently support that  $p$ , but we might not believe that  $p$  in the way in which evidentialism requires us to do. This happens, for instance, when we believe that  $p$  for reasons that have nothing to do with the fact that our evidence supports  $p$ .<sup>4</sup> Sometimes these mistakes are inevitable given our cognitive limitations. Sometimes they are due to bad luck. Other times, instead, they are attributable to our negligence. Either way, just as appropriately believing on the basis of one's supporting evidence is thought to be instrumentally valuable insofar as it maximizes the chances of having true beliefs, failing to do so seems to be instrumentally problematic insofar as it generally obstructs our gaining true beliefs.

Consider now *epistemic paternalism*. Epistemic paternalism is traditionally conceived of as the practice of (1) limiting someone's freedom, (2) without their consent, (3) in order to improve someone's epistemic situation.<sup>5,6</sup> Epistemic paternalism has traditionally been developed with two assumptions in mind. First, it has been developed with the assumption that epistemic paternalistic practices are meant to improve the subject's epistemic situation. Second, it has also traditionally been built against the background of veritism. In other words, epistemic paternalistic practices are traditionally considered to be successful insofar as, by limiting a subject B's freedom, they maximize B's ratio of true

4. Epistemologists talk about doxastic as opposed to propositional justification. Propositional justification is usually understood as a property of a proposition  $p$  given one's evidence; doxastic justification is a property of a *belief* given one's evidence. The case in which my evidence supports  $p$ , but I believe that  $p$  for reasons that have nothing to do with my evidence, is a case in which I have propositional justification, but I lack doxastic justification.

5. See Kristoffer Ahlstrom-Vij, *Epistemic Paternalism: A Defence* (London: Palgrave MacMillan, 2013), section 2; Duncan Pritchard, "Epistemic Paternalism and Epistemic Value," *Philosophical Inquiries* 1, no. 2 (2013): 10; Elizabeth Jackson, "Epistemic Paternalism, Epistemic Permissivism, and Standpoint Epistemology," in *Epistemic Paternalism Reconsidered: Conceptions, Justifications, and Implications*, ed. Amiel Bernal and Guy Axtell (Lanham, MD: Rowman and Littlefield, 2020), 201; and Emma C. Bullock, "Knowing and Not-Knowing for Your Own Good: The Limits of Epistemic Paternalism," *Journal of Applied Philosophy* 35, no. 2 (2018): 434.

6. Note that the idea of epistemic paternalism is a special case of the more liberal kind of paternalism, e.g., nudges, endorsed by Richard H. Thaler and Cass R. Sunstein, *Nudge: Improving Decisions about Health, Wealth, and Happiness* (New York: Penguin, 2003) in the political sphere. I consider cases of epistemic nudging in the section titled "Weak Epistemic Paternalism."

to false beliefs.<sup>7</sup> The exclusionary rule that classifies character evidence to be inadmissible in court is often thought of as being a paradigmatic instance of epistemic paternalism. This rule is generally motivated by the fact that jurors are likely to overestimate its relevance and, therefore, to be misled by it.<sup>8</sup>

As both evidentialism and epistemic paternalism have been traditionally developed against the background of veritism, we might wonder whether the debates on these two topics can fruitfully complement each other. After all, evidentialism is concerned with how to get to the truth; epistemic paternalism is concerned with how we can help others get to the truth. This paper explores this connection by asking whether and to what extent epistemic paternalistic practices can help the subject of the intervention to overcome some common and important evidential mistakes in a satisfying way. In this paper, I challenge the veritist assumption behind traditional accounts of epistemic paternalism. I do so by considering both strong epistemic paternalistic practices, as well as weak epistemic paternalistic practices, e.g., epistemic nudges. Finally, I show the implications that these theoretical considerations have for education. Before getting into the details of the examples, however, in the next section I clarify the scope and the questions I address in this paper.

#### CLARIFYING THE SCOPE AND QUESTIONS

Before considering examples of epistemic paternalistic practices, two points are worth making.

First, much of the literature on epistemic paternalism so far has been concerned with the following question: Are epistemic paternalistic strategies justified? However, as Elizabeth Jackson recently stresses, we need to clarify what we mean by *justified* here.<sup>9</sup> A common way to think about whether epistemic paternalistic interventions are justified is to ask whether they are *all things considered* justified, and not merely justified with respect to whether they benefit the subject of the intervention *epistemically*. This brings up another important issue: if epistemic paternalistic practices limit the subject's freedom and autonomy, can they *ever* be all-things-considered justified? Given the ethical imperative of respecting people's freedom and autonomy, the defender of epistemic paternalistic practices needs to provide us with a convincing story as to why

7. Alvin I. Goldman, "Epistemic Paternalism: Communication Control in Law and Society," *Journal of Philosophy* 88, no. 3 (1991): 124; and Ahlstrom-Vij, *Epistemic Paternalism*, 51.

8. Goldman, "Epistemic Paternalism," 117–118. Note that this definition of epistemic paternalism doesn't specify whether the subject whose freedom is limited should be the same as the subject whose situation is improved by the intervention. In this case, the jurors are both the people whose freedom is limited and whose *epistemic* situation is improved. However, it's plausible to say that this intervention is meant to promote the overall benefit of someone other than the jurors, e.g., the defendant. In the cases I consider in this paper the subject whose freedom is limited is also the subject whose epistemic situation is supposed to be improved by the intervention. Thanks to Jesper Kallestrup for raising this point to me in conversation.

9. Elizabeth Jackson, "What's Epistemic about Epistemic Paternalism," in *Epistemic Autonomy*, ed. Jonathan Matheson and Kirk Lougheed (New York: Routledge, 2022).

acting in a way that promotes someone else's epistemic good is worth jeopardizing their freedom and autonomy.<sup>10,11</sup> In this paper, I will bracket the question about the all-things-considered justification, and assume that the paternalistic interventions that I consider do not limit people's autonomy and freedom in any morally problematic way.

Second, while most of the discussion in the literature has focused on the question of whether epistemic paternalism is all-things-considered justified, on whether epistemic paternalism is a genuinely *epistemic* kind of paternalism, or on who is entitled to apply epistemic paternalistic interventions, there has not been much philosophical discussion on whether and to what extent we should expect these practices to be *effective* in making the subject of the intervention a better epistemic agent.<sup>12</sup> While a complete and satisfying answer to this question will ultimately rest at least partly on empirical findings, considering the epistemological discussion on evidence, knowledge, and virtues can give us theoretical reasons to expect some strong paternalistic practices to be effective in allowing the subjects of the intervention to overcome important evidential mistakes, while remaining unsatisfying tools to make them better epistemic agents overall. However, later I will argue that we should not be overly pessimistic about the prospects of using weak epistemic paternalistic techniques to make the subject of the intervention a better epistemic agent. Finally, I will turn to the practical implications of these theoretical considerations.

#### STRONG EPISTEMIC PATERNALISM

##### EPISTEMIC PATERNALISM, EVIDENTIALISM, AND TRUE BELIEF

Can limiting B's evidence allow B to overcome evidential mistakes? Consider the following case.

##### CONFORMITY TO EVIDENCE

Peter is scared of spiders and, as a consequence, he always believes that there are no spiders around. Despite being in the presence of spider webs, Peter believes that there are no spiders around. Jane, who knows Peter very well, and who cares deeply about Peter's epistemic and rational well-being, decides to change the environment around him, e.g., by dusting all spider

10. See, for example, Pritchard, "Epistemic Paternalism and Epistemic Value"; and Bullock, "Knowing and Not-Knowing for Your Own Good."

11. Note that nudging is generally thought of as being compatible with respecting the autonomy of the subject of the intervention. I will consider nudging in the section titled "Weak Epistemic Paternalism."

12. Emma Bullock and Elizabeth Jackson have been concerned with the former project (see Bullock, "Knowing and Not-Knowing for Your Own Good"; and Jackson, "What's Epistemic about Epistemic Paternalism," section 7.2). Michel Croce has been concerned with the latter project (see Michel Croce, "Epistemic Paternalism and the Service Conception of Epistemic Authority," *Metaphilosophy* 49, no. 3 [2018]: 305–327). Daniella Meehan is the only philosopher who, to my knowledge, has been concerned with whether these practices are effective (see Daniella Meehan, "Epistemic Vice and Epistemic Nudging: A Solution?," in *Epistemic Paternalism: Conceptions, Justifications and Implications*, ed. Guy Axtell and Amiel Bernal [London: Rowman & Littlefield, 2020], 249–261). I consider Meehan's work in the section titled "Weak Epistemic Paternalism."

webs in the house. Peter can't see any spider webs around anymore. He continues to believe that there are no spiders around.

Before Jane's intervention, Peter commits important evidentialist mistakes: despite his evidence (which includes the proposition that there are spider webs), he fails to believe what his evidence supports (e.g., that there are spiders around), and instead believes a proposition that is not supported by the evidence (i.e., that there are no spiders around). Jane's intervention seems to improve Peter's epistemic situation. After Jane's intervention, Peter's belief that there are no spiders around is supported by his total evidence, which now plausibly includes the true proposition that there are no spider webs around. And yet, the way in which Peter is forming beliefs about spiders is still problematic. An obvious way in which Jane's intervention is unsuccessful is that it does not allow Peter to form true beliefs. But what if Jane also gets rid of the spiders around? In this modified case, Jane's intervention would count as being successful from a veritist point of view. Jane is modifying Peter's environment in a way that guarantees that Peter's evidence supports whatever Peter believes, and in a way that makes Peter's beliefs true.

One might conclude that epistemic paternalistic interventions like the one above are thus successful. However, there is something unsatisfying about Jane's intervention. To begin with, Peter is still committing an important evidentialist mistake: he does not believe that *p* on the basis of his supporting evidence for *p*.

Consider now the following case.

#### EVIDENCE-SUPPORT ASSESSMENT

Jenny knows that she should believe things on the basis of her evidence, and not according to a whim. However, she often makes mistakes as to what her evidence actually supports. In particular, she tends to believe whatever she reads in a book, regardless of whether it is a fiction or a history book. As a consequence, when she comes across a fantasy book, she believes false propositions, e.g., that fairies and unicorns are real. Believing on the basis of evidence is thus not always a means to true belief. To maximize the chances of her forming true beliefs, Jenny's parents leave her with only one fantasy book and all history books. Jenny randomly picks up a book, which happens to be a history book, which says that World War 2 ended in 1945 (*p*). She believes that *p* on the basis of her evidence that *p*.

Contrary to Peter, Jenny is aware that she should believe what her evidence supports. Furthermore, while Jane modifies Peter's environment in a way that, in turn, affects what evidence Peter has, Jenny's parents directly limit the amount of evidence that Jenny has. By limiting the amount of evidence that Jenny would be likely to misinterpret, Jenny's parents maximize the chances of her gaining true beliefs; moreover, their intervention allows her to do so *by appropriately believing on the basis of supporting evidence*. One could take this example to conclude that strong epistemic paternalistic practices that involve limiting the evidence the subject of the intervention has can be effective: they can allow the subject of the intervention to fix important evidential mistakes, believe in the way prescribed by evidentialism, and, by doing so, achieve true beliefs.

#### EPISTEMIC PATERNALISM, EVIDENTIALISM, AND KNOWLEDGE

Let us take a step back and ask, Why is basing one's belief on the supporting evidence important? Remember that, as mentioned above, evidentialism is

traditionally motivated by appealing to veritism. But if believing on the basis of the evidence is only instrumentally valuable because it is a means to true beliefs, then it is not clear why there is any problem with Peter's case. After all, once Jane has eliminated all the spiders around, Peter's beliefs are true, despite not being based on his supporting evidence.

One reaction one might have from considering cases like Peter's is to say that evidentialism should be motivated by appealing to Gnosticism (as opposed to veritism). Take Gnosticism to be the view on which *knowledge* (as opposed to true belief) is the fundamental epistemic good we should care about.<sup>13</sup> Evidential mistakes should thus be avoided insofar as they can obstruct the acquisition of knowledge. While philosophers have mainly assumed veritism when discussing strong epistemic paternalism, cases like Peter's seem to suggest the strong epistemic paternalistic practices that we should implement are those that aim at producing *knowledge* in the subject of the intervention (rather than mere true belief).

Spelling out exactly what knowledge-centered epistemic paternalistic practices look like is something I will have to do in another paper. However, for our purposes, the following consideration will be enough. If one takes strong epistemic paternalistic practices to be instrumentally valuable insofar as they allow the subject of the intervention to acquire knowledge, then many epistemologists would say that the way in which Jenny's parents limit her evidence is unsuccessful. For instance, many epistemologists would claim that Jenny's belief is not safe: she could have easily formed a false belief by picking up the only fantasy book in the room. Given that they would take safety to be necessary for knowledge, they might point out that, despite having a true belief based on supporting evidence, Jenny's belief fails to constitute *knowledge*.<sup>14</sup>

Could the problem be that the intervention implemented by Jenny's parents is *not strong* enough? Imagine that Jenny's parents decide to leave her *only* with history books. Jenny picks up a book and believes that *p* on the basis of what the book says. In this case, it seems plausible to ascribe knowledge of the historical truths to Jenny. Her belief is true, based on supporting evidence, and could not have easily been mistaken. Limiting evidence in this way seems to allow Jenny to avoid committing evidential mistakes and, by doing so, to acquire knowledge.

#### EPISTEMIC PATERNALISM, EVIDENTIALISM, AND VIRTUES

In the above case, strongly limiting Jenny's evidence seems to improve her epistemic situation. After her parents' intervention, Jenny avoids making evidential

13. Timothy Williamson, *Knowledge and Its Limits* (Oxford: Oxford University Press, 2000); and Clayde Littlejohn, "La verdad en el gnosticismo" [The Truth in Gnosticism], *Análisis. Revista de investigación filosófica* 3, no. 2 (2016): [https://doi.org/10.26754/ojs\\_arif/a.rif.201621568](https://doi.org/10.26754/ojs_arif/a.rif.201621568).

14. For the claim that knowledge requires safety, see Williamson, *Knowledge and Its Limits*, 147; Duncan Pritchard, *Epistemic Luck* (Oxford: Oxford University Press, 2005); and Ernest Sosa, "How to Defeat Opposition to Moore," *Philosophical Perspectives* 13 (1999): 141–153.



mistakes (i.e., wrongly assessing what her evidence supports) and, as a consequence, she acquires knowledge.

Is the assessment of this case satisfying? One possible reason to find these kinds of interventions unsatisfying is that, despite the fact that they allow Jenny to appropriately believe on the basis of her supporting evidence in that occasion, such interventions do not seem to tackle her *overall attitude* toward how careful she should be in assessing what her evidence supports. Successful strong epistemic paternalistic practices will thus be the ones that make a difference in whether Jenny's believing on the basis of her supporting evidence is the manifestation of some epistemic virtues.

Let me first distinguish between two ways of understanding epistemic virtues. Let us start with virtue reliabilists. These are concerned with reliable cognitive faculties, or "faculty-virtues," and include, e.g., perception and memory. To put it in Heather Battaly's terms, "Virtue reliabilists conceive of virtues as *instrumentally* valuable."<sup>15</sup> If these virtues are seen as instrumentally valuable, what are they valuable for? One way to answer this question is to say that these faculty-virtues are important because they reliably generate true belief.<sup>16</sup> But if truth is what is valuable, then we are forced to reassess both Peter's case and the original issue in Jenny's case: systematically implementing epistemic paternalistic practices that maximize truth acquisition, e.g., in Peter's and Jenny's cases, are perfectly satisfying interventions.

A better answer to the above question is that faculty-virtues are valuable insofar as they lead to knowledge.<sup>17</sup> One could thus implement evidentialism with virtue reliabilism and say that basing one's belief on one's supporting evidence — as evidentialism requires us to do — is a means to knowledge only if it is the result of some reliable epistemic virtue.<sup>18</sup>

Can appealing to virtue reliabilism help explain what is unsatisfying about the epistemic paternalistic strategy implemented by Jenny's parents? One could say that, in order to count as *knowledge*, Jenny's believing on the basis of her evidence would have to be the manifestation of some epistemic virtue. Her true belief would have to be creditable to some of her reliable cognitive functions. However, the relevant cognitive faculties Jenny employs in the scenario seem to be working pretty well. Jenny is not believing historical truths based on a whim. Instead, she believes those truths *because* a history book says so. If Jenny were to take a history

15. Heather Battaly, "Varieties of Epistemic Vices," in *The Ethics of Belief*, ed. Jonathan Matheson and Rico Vitz (Oxford: Oxford University Press, 2014), 54–55 (emphasis added).

16. See, for example, John Greco, "Virtues in Epistemology," in *Oxford Handbook of Epistemology*, ed. Paul Moser (New York: Oxford University Press, 2002).

17. Lisa Miracchi, "Competence to Know," *Philosophical Studies* 172, no. 1 (2015): 29–56; and Christoph Kelp, *Good Thinking: A Knowledge-First Virtue Epistemology* (New York: Routledge, 2019).

18. Lisa Miracchi, "When Evidence Isn't Enough: Suspension, Evidentialism, and Knowledge-First Virtue Epistemology," *Episteme* 16, no. 4 (2019): 413–437.

exam, she would answer all the questions correctly, and it would thus be strange to deny her knowledge of those historical facts. Retaining information about history on the basis of a history book seems to be a perfectly good way of gaining knowledge of historical events. On the other hand, one might insist that we should identify the method of belief formation differently. The method Jenny is employing is not the method of believing what history books say, but the method of believing whatever books she happens to come across say. If this is so, then Jenny lacks knowledge and that is what makes the intervention unsatisfying.<sup>19</sup>

According to this view, the question of whether limiting B's evidence is an effective way for B to acquire knowledge is ultimately going to depend on how we identify the target method of belief formation, something that is very hard to do.<sup>20</sup> If we want to find theoretical reasons to think epistemic paternalistic practices can be successful in allowing B to overcome evidential mistakes, in a way that allows B to achieve knowledge, we need to look elsewhere.

Let us now consider virtue responsibilism. Virtue responsibilists understand epistemic virtues as "character traits," dispositions, or "trait-faculties." Examples of epistemically virtuous character traits are open-mindedness, epistemic humility, thoroughness, and critical thinking. Epistemic virtues understood as character traits are usually considered in opposition to what Quassim Cassam calls epistemic vices: character traits that systematically or normally obstruct the attainment of knowledge-conducive inquiry.<sup>21</sup> Examples of epistemic vices include dogmatism, "intellectual pride, negligence, idleness, cowardice, conformity, carelessness, rigidity, prejudice, wishful thinking, closed-mindedness, insensitivity to detail, obtuseness, and lack of thoroughness."<sup>22</sup>

One might try to implement evidentialism with virtue responsibilism and say that not only should one believe on the basis of one's supporting evidence, but that this basing should be the result of having a *general disposition to be sensitive to the evidence one has*. That is, it should be an instance of a more general disposition to appropriately believe on the basis of one's supporting evidence. While focusing on faculty-virtues does not explain why the epistemic paternalistic practices of Jenny's parents are unsatisfying, considering Jenny's general disposition to assess her evidence in a certain way might do the explanatory job. A successful strong epistemic paternalistic practice should be able to mitigate her general gullibility. And yet, despite the fact that her parents' intervention allows Jenny to believe that

19. Then, again, if we identify the method with "believing any book I find in the room," we should conclude that this method is reliable and Jenny does have knowledge.

20. Consider the "generality problem" traditionally raised for process reliabilism. See Earl Conee and Richard Feldman, "The Generality Problem for Reliabilism," *Philosophical Studies* 89, no. 1 (1998): 1–29.

21. Quassim Cassam, "Vice Epistemology," *The Monist* 99, no. 2 (2016): 159–180; and Quassim Cassam, *Vices of the Mind: From the Intellectual to the Political* (Oxford: Oxford University Press, 2021).

22. Linda Trinkaus Zagzebski, *Virtues of the Mind* (Cambridge, UK: Cambridge University Press, 1996), 152.



$p$  on the basis of her supporting evidence, it does not seem to change or mitigate Jenny's broader gullibility.

One might ask why we should care about Jenny's general gullibility. One idea is that epistemically virtuous character traits are instrumentally useful insofar as they are required for obtaining knowledge.

Crucially, I believe that the idea that knowledge requires some virtuous epistemic character traits can be questioned. First, consider cases of perceptual knowledge. Someone who is generally full of vicious character traits can nevertheless know that she has hands by looking at her hands. Second, one might insist that knowledge at least requires those character virtue traits that are relevant for assessing evidence correctly. But consider Saul Kripke's dogmatist paradox: in cases in which I know that  $p$ , I can also come to know that any future evidence against  $p$  will be misleading. Therefore, when I know that  $p$  and I am presented with misleading evidence against  $p$ , remaining dogmatic, e.g., by avoiding or ignoring evidence against  $p$ , seems to be an effective way of preserving my knowledge. One way to read Kripke's paradox is to take it as providing us with an example of how character traits traditionally considered to be "epistemically vicious" (e.g., dogmatism) are not always incompatible with knowledge; rather, they can allow us to preserve our knowledge.<sup>23</sup> Note that, according to Cassam, this is not the conclusion we should draw from Kripke's paradox. Cassam argues that the cases in which one knows that  $p$  and decides, dogmatically, to avoid or ignore any future evidence against  $p$  are generally cases in which one lacks epistemic self-confidence. And yet, "lack of epistemic self-confidence is itself a threat to knowledge."<sup>24</sup> Therefore, Cassam says, dogmatic attitudes *do* systematically obstruct the acquisition of knowledge after all. I believe that Cassam's analysis rests on an ambiguity over what it means to lack epistemic self-confidence. On one reading, the subject who decides at a time  $t$  to avoid or ignore future evidence against  $p$ , lacks, at  $t$ , confidence in the proposition that  $p$ . On a different reading (presumably the one he has in mind), the subject who decides at a time  $t$  to avoid or ignore future evidence against  $p$  lacks, at  $t$ , the confidence in her ability to refute potential future misleading evidence. But note that while the former might be incompatible with  $S$  having a full belief (and knowledge) that  $p$  at  $t$ , it is not clear why lacking self-confidence in the latter sense is incompatible with knowledge at  $t$ . However, once we grant that dogmatic attitudes are compatible with knowledge, there is no reason to think that other paradigmatic examples of epistemic vices will be incompatible with the attainment of knowledge.<sup>25</sup>

23. See Peter Millican, "The Riddle of Hume's Treatise: Skepticism, Naturalism, and Irreligion," *British Journal for the History of Philosophy* 19, no. 2 (2011): 348–353; Cassam, *Vices of the Mind*, chap. 5; and Saul A. Kripke, "On Two Paradoxes of Knowledge," in *Philosophical Troubles: Collected Papers*, vol. 1, ed. Saul A. Kripke (Oxford: Oxford University Press, 2011).

24. Cassam, *Vices of the Mind*, 102–110.

25. Similarly, in various papers Maria Lasonen-Aarnio has argued for cases of "unreasonable knowledge," namely, cases in which one has knowledge despite manifesting a general bad epistemic disposition. See Maria Lasonen-Aarnio, "Unreasonable Knowledge," *Philosophical Perspectives* 24, no. 1 (2010): 1–21.

Finally, and related to this point, even if one is not convinced by the idea that epistemic vices are compatible with knowledge, in order to defend the idea that virtue traits are required for knowledge, one needs to define *which* character traits are allegedly required for knowledge, and which ones are not. Is open-mindedness required for knowledge?<sup>26</sup> If anything, the problem with Jenny's gullibility seems to be that she is *too* open-minded. Perhaps knowledge requires *critical scrutiny*? This might make sense of Jenny's case, but it is hard to believe that critical scrutiny is *always* required for knowledge. Being *too critical* of one's available evidence can sometimes constitute an epistemic vice.<sup>27</sup>

Given these considerations, it seems appropriate to ascribe knowledge to Jenny: her being gullible just does not seem to interfere with her achieving knowledge in this occasion.

So why should we care about her gullibility if knowledge is compatible with being gullible? One plausible answer is to say that, although knowledge does not require having a *general disposition* to correctly assess one's evidence, such a disposition is important because it is *generally* or *normally* conducive to knowledge. More generally, one might think that good dispositions to appropriately select, assess, and use one's evidence when believing things generally do underpin the broader skill of "thinking critically," which, one might claim, is what we should ultimately care about.

Determining why we should focus on epistemically virtuous character traits goes beyond the scope of this paper. However, what matters for our purposes is to note that, while limiting the evidence available to someone can be an effective way to modify their external environment and allow them to avoid making evidential mistakes, and perhaps even to achieve knowledge, it is difficult to see how these interventions can have an effect on the broader epistemic character traits of the subject of the intervention.

#### WEAK EPISTEMIC PATERNALISM

In the previous sections, I have considered strong epistemic paternalistic practices, and I have argued that, whether they are motivated by gnosticism or veritism, there are theoretical reasons to think they will not be effective in overcoming evidential *vices*. But what about so-called "weak epistemic paternalism," e.g., epistemic nudging? Can epistemic nudging help with overcoming epistemically vicious traits? In this section, I consider Daniella Meehan's argument for pessimism.<sup>28</sup> Showing that nudging techniques can be effective to tackle epistemic vices is an empirical matter, something I will not try to establish here. My aim is more modest. In what follows, I will provide theoretical reasons to believe

26. See Cassam, *Vices of the Mind*, 105.

27. See C. Thi Nguyen, "Echo Chambers and Epistemic Bubbles," *Episteme* 17, no. 2 (2020), 156.

28. Meehan, "Epistemic Vice and Epistemic Nudging."

Meehan's pessimism is unwarranted. By contrast, recent empirical findings suggest that we can be more optimistic about the prospects of implementing nudging techniques to help students eradicate or mitigate evidential vices.

#### HOW STRONG IS THE CASE AGAINST EPISTEMIC NUDGING?

The concept of nudging, as developed by Cass Sunstein and Richard Thaler, refers to the way in which we can positively influence people's choices and actions simply by presenting options and information to them in certain ways.<sup>29</sup> A paradigmatic example of nudging is that of showing healthy snacks at eye level in vending machines so that people are more likely to pick the healthy option. As with epistemic paternalism more generally, there is no consensus on whether nudging techniques are always morally justified.<sup>30</sup> However, most would consider them to be less morally problematic than strong epistemic paternalistic practices for the following reason: our choices and behavior are often irrationally and unconsciously influenced by how things and options are presented to us. Nudging techniques exploit people's irrational behavior to their advantage: they manipulate the environment in a way that influences people to do what is best for them. Furthermore, nudging techniques are often thought of as examples of "weak" paternalistic interventions, or libertarian paternalism, because, to put it in Thaler and Sunstein's words,

[Epistemic nudges] alter people's behavior in a predictable way without forbidding any options or significantly changing their economic incentives. To count as a mere nudge, the intervention must be easy and cheap to avoid. Nudges are not mandates. Putting fruit at eye level counts as a nudge. Banning junk food does not.<sup>31</sup>

Similarly to what I did in the first half of this paper, I am going to assume that nudges are not *morally* problematic. Instead, here I will be concerned with the question of whether they are justified from an instrumentalist point of view. More precisely, here I consider Meehan's recent pessimism and argue that there are *theoretical* reasons to believe that her case against the use of nudges to overcome epistemic vices has yet to be established.

Borrowing the terminology from the literature on dispositions, Meehan argues that nudging techniques can at most *mask* epistemic vices. These, however, are going to manifest themselves again once the intervention is over.<sup>32</sup> In particular, she argues that cases of epistemic nudging *seem* to mitigate epistemic vices only if we are working with a "shallow" (and, according to her, mistaken) conception of epistemic vice, one that identifies the epistemic vice "within the practices in which it is typically manifested or the projects of inquiry it obstructs."<sup>33</sup> According

29. Thaler and Sunstein, *Nudge*.

30. For an overview, see Andreas T. Schmidt and Bart Engelen, "The Ethics of Nudging: An Overview," *Philosophy Compass* 15, no. 4 (2020), <https://doi.org/10.1111/phc3.12658>.

31. Thaler and Sunstein, *Nudge*, 6.

32. Meehan, "Epistemic Vice and Epistemic Nudging."

33. *Ibid.*, 255.

to Meehan, once we correctly identify epistemic vices as “deep psychological disposition[s] ... influenced by societal structures,” then we realize that epistemic nudges (allegedly) do not mitigate the dispositional trait; they merely *mask* it. To put it in Meehan’s words,

[Epistemic Nudging (EN)] does not change this vice in any way, just like the bubble wrap [does not] change the fragility of the vase, but only masks it, and when EN practices are not employed the vice is still present, just like how the fragility of the vase still remains when the bubble-wrap is removed.<sup>34</sup>

Furthermore, according to Meehan, not only are nudging techniques not effective, but they are actually counterproductive, insofar as they lead to the vice of epistemic laziness. Meehan’s argument runs schematically as follows. First, epistemic laziness arises when subjects do not employ their rational and reflective epistemic capacities. Second, drawing on Evan Riley’s account of epistemic nudges, Meehan takes EN as altering “our epistemic capacities to reason and think critically.”<sup>35</sup> Therefore, the thought goes, epistemic nudges can often cause epistemic laziness in the long run:

If decisions are made for us or we are pushed to make certain decisions, despite being for the best (e.g., being less dogmatic), the violation of autonomy give way to new vices, as argued, specifically the vices of epistemic laziness. ... Like how muscles are lost over time if they are not exercised, epistemic capacities that are not exercised due to the practice of EN will also be lost, leading to the creation of epistemic laziness.<sup>36</sup>

I believe Meehan is correct in saying that paternalistic interventions will not be effective in tackling epistemic vices unless they act at the *dispositional* level. However, as things stand, Meehan’s pessimism about nudging techniques is unjustified. First, consider again Meehan’s claim that epistemic nudges are not effective at the *dispositional* level. As Meehan says, given the dispositional nature of epistemic vices, EN can at best *mask* the target epistemic vice, which will manifest again once the intervention is over. Epistemic nudges thus cannot eradicate or improve epistemic vices. But note that whether the effect of nudging remains after the intervention is over is an empirical matter over which no consensus has been achieved yet.<sup>37</sup> For instance, according to a recent study, whether the nudges can have a temporal spillover effect depends on the kinds of behavior in question.<sup>38</sup> Furthermore, note that if EN were *ineffective* at the dispositional level — as Meehan thinks they are — then it is not clear why we should expect them to *generate* new negative dispositions, such as epistemic

34. Ibid.

35. Ibid., 257.

36. Ibid., 256

37. This phenomenon is often known as the “Temporal Spillover Effect.”

38. Merije Van Rookhuijzen, Emely De Vet, and Marieke A. Adriaanse, “The Effects of Nudges: One-Shot Only? Exploring the Temporal Spillover Effects of a Default Nudge,” *Frontiers in Psychology* 12 (2021): <https://doi.org/10.3389/fpsyg.2021.683262>.

laziness. The other side of the coin is that, if it is possible for EN to bring about a *negative* change at the dispositional level, then we cannot in principle rule out the possibility that they can also affect someone's epistemic dispositions *positively*.

Finally, note that, as mentioned above, Meehan's argument (allegedly) showing that nudging techniques lead to epistemic laziness rests on a specific conception of EN, one defended by Riley.<sup>39</sup> According to Riley, EN bypasses rational reflection in a way that hinders our epistemic capacities to reason and think critically. But which epistemic capacities are relevant for critical thinking? Is the ability to believe on the basis of one's supporting evidence the target epistemic capacity we should be concerned with? Or is it the ability to gather enough and appropriate evidence? Furthermore, even if we had a satisfying answer to this question, note that Riley himself acknowledges that "some nudges present nudges with truths, plausible claims, or information of which they were not previously aware, and thereby get them engaged in relevant, practically oriented, rational inference."<sup>40</sup> Riley believes that this effect does not suffice to always make EN *morally* permissible. But what matters for our purposes is not whether EN is morally justified, but rather whether it can be effective in overcoming some epistemic vices. Even if EN sometimes fails to directly engage a person's full rational capacities, we should not conclude that EN are always ineffective when it comes to mitigating vices of rationality. While these considerations do not show that nudging techniques would be effective in mitigating the evidential vices of the subject of the intervention, they do suggest that the case against them needs to be yet established.

### CONSEQUENCES FOR EDUCATION

Can epistemic paternalistic practices be used in an educational context to help students become better epistemic agents? A complete and satisfying answer will ultimately depend at least in part on empirical findings. Furthermore, what counts as successful will in part rest on what one takes the ultimate aim of education to be. However, in this paper I have shown that looking at the epistemological discussion on the relation between knowledge, evidence, and epistemic virtues can give us theoretical reasons to determine which epistemic paternalistic practices we should expect to be effective in fixing someone's evidential mistakes, fostering knowledge, and making someone a better epistemic agent more generally. Let me conclude this paper by showing some of the consequences that the above theoretical considerations have for education.

*Minimizing evidential mistakes and maximizing true belief.* Strong epistemic paternalistic strategies that involve modifying the student's environment or limiting the amount of evidence the student has access to may be effective in fixing some evidential mistakes in a way that allows the student to achieve true beliefs.

39. Evan Riley, "The Beneficent Nudge Program and Epistemic Injustice," *Ethical Theory and Moral Practice* 20, no. 3 (2017): 597–616.

40. *Ibid.*, 600.

*Minimizing evidential mistakes and maximizing knowledge.* Whether strong epistemic paternalistic strategies will be successful in fostering knowledge in the subject of the intervention will depend on which account of knowledge turns out to be true. On the one hand, if knowledge requires an externalist component like the safety of the belief or the reliability of the belief-forming method, then the prospects of strong epistemic paternalistic interventions that involve modifying the students' environment or limiting the students' evidence are promising. However, many epistemologists would also claim that whether a belief is safe and whether it is the result of a reliable faculty-virtue depends on how the method of belief-formation is identified. This is a notoriously difficult thing to determine, especially given that the method of belief formation needs to be specific to the target student. As things stand, it is thus difficult to establish in advance whether strong epistemic paternalistic strategies can be effective and reliable tools to use in the classroom to foster the acquisition of knowledge in students.

*Minimizing evidential mistakes and improving virtuous character traits.* The cases presented in this paper suggest that maximizing the amount of true beliefs or knowledge the student has is not enough for making the student a better epistemic agent. Instead, a better epistemic agent is one who has general virtuous character traits that translate into a general disposition to appropriately select, assess, and use one's evidence to believe things. The problem is that it is not clear how merely limiting the amount of information the student has access to can affect the student at the dispositional level. However, I have argued that knowledge is compatible with vicious character traits. This theoretical consideration has important practical implications: for once the link between epistemic vices/epistemic virtues and ignorance/knowledge is severed, then we should not expect paternalistic interventions that aim at maximizing the chances that students acquire *knowledge* to be effective in allowing the students to overcome their epistemic vices too. Practically, this means that we might need two different pedagogies to achieve these two distinct goals.

Can weak epistemic paternalistic practices, e.g., nudging, be effective in making someone a better epistemic agent? In the previous section, I argued that Meehan's case against epistemic nudges is unjustified. Whether we should think that epistemic nudges are effective is of course an empirical matter that I do not try to establish here. However, let me conclude with one consideration. As mentioned above, one worry Meheen has is that nudges lead to epistemic laziness. However, recent empirical findings do not justify this outright pessimism. For instance, drawing on a distinction made by Pelle Guldberg Hansen and Andreas Maaløe Jespersen, Robert Weijers and colleagues have recently considered two types of nudging techniques.<sup>41</sup> Both techniques engage with people's automatic

41. Pelle Guldberg Hansen and Andreas Maaløe Jespersen, "Nudge and the Manipulation of Choice: A Framework for the Responsible Use of the Nudge Approach to Behaviour Change in Public Policy," *European Journal of Risk Regulation* 4, no. 1 (2013): 3–28; and Robert J. Weijers, Björn B. de Koning, and Fred Paas, "Nudging in Education: From Theory towards Guidelines for Successful Implementation," *European Journal of Psychology of Education* 36, no. 3 (2021): 883–902.



system, i.e., what Daniel Kahneman has called system 1.<sup>42</sup> However, while Type 1 nudging engages with the automatic system without involving any kind of reflective thinking, Type 2 nudging “engages with the automatic system in order to trigger reflective thinking that subsequently shapes behaviour.”<sup>43</sup> Engraving an image of a fly in the urinal to reduce spillage is taken to be an example of Type 2 nudging. The image of the fly engages the automatic system, which in turn triggers a reflective response: the user starts actively paying attention. But if Type 2 nudging is a genuine possibility, then even if epistemic laziness allegedly arises when a subject does not employ her rational and reflective capacities, Meehan’s claim that nudging techniques bring about epistemic laziness seems unwarranted. By contrast, using Type 2 nudging techniques in an educational context might be a promising way to go. Defining what these epistemic nudges could be like in an educational context goes beyond the scope of this paper. But we can imagine what they could be like. They would not be interventions that change the world or the student’s evidence in a way that conforms to their beliefs. Instead, they would be nudges that trigger the student’s reflective capacity to ask themselves *why* and *how* they believe what they do.

---

42. Daniel Kahneman, *Thinking, Fast and Slow* (New York: Farrar, Straus and Giroux, 2011).

43. Weijers et al., “Nudging in Education,” 889.

---

I AM GRATEFUL to Adam Carter, Edoardo Cavašin, Jesper Kallestrup, Joshua Thorpe, and an anonymous referee for useful comments on this paper. This work for this paper has been supported by the AHRC grant number AH/T002638/1, 758539.