

# Disagreement and the value of self-trust

Robert Pasnau

Published online: 4 December 2014  
© Springer Science+Business Media Dordrecht 2014

**Abstract** Controversy over the epistemology of disagreement endures because there is an unnoticed factor at work: the intrinsic value we give to self-trust. Even if there are many instances of disagreement where, from a strictly epistemic or rational point of view, we ought to suspend belief, there are other values at work that influence our all-things considered judgments about what we ought to believe. Hence those who would give equal-weight to both sides in many cases of disagreement may be right, from the perspective of pure rationality. But their critics are right too, in seeing something undesirable in the consequences of giving equal weight. Among epistemologists, there is a tendency to set aside trust and other such non-epistemic factors, on the grounds that these are not germane to their topic. But ultimately, the value of self-trust shows signs of encroaching on the strictly epistemic question of when our beliefs may be said to be justified. Hence again, even if the equal-weighters are right about what is rational, they may be wrong about what knowledge requires.

**Keywords** Disagreement · Knowledge · Alienation · Self-trust

## 1 Introduction

I say yes. You say no. I feel that my reasons are strong, but I also feel that you are as smart and well informed as I am. I know your reasons, and do not find them persuasive, but I know that you also know my reasons, and that you in turn do not find them persuasive. After reflection upon all of this, I revise my belief. I say I don't know.

---

R. Pasnau (✉)  
University of Colorado, Boulder, CO, USA  
e-mail: pasnau@colorado.edu

With predictable irony, the debate over how best to handle such disagreements seems far from yielding consensus. On one side stand those who stress the apparently equal epistemic credentials of one's opponents, and the rationality of giving their views consideration equal to one's own. On the other side are those who focus on the apparently intolerable consequences of such epistemic impartiality: the pressure to suspend judgment across a wide range of topics, and the implication that, if we do not suspend judgment, then our beliefs will not count as knowledge. I propose that the reason this debate remains unsettled is that there is an unnoticed factor at work: the intrinsic value we give to self-trust. Even if there are many instances of disagreement where, from a strictly epistemic or rational point of view, we ought to suspend belief, there are other values at work that influence our all-things-considered judgments about what we ought to believe. Hence the proponents of impartiality may be right, from the perspective of pure rationality. But their critics are right too, in seeing something undesirable in the consequences of being impartial.

Among epistemologists, there is a tendency to set aside trust and other such non-epistemic factors, on the grounds that these are not germane to their topic. But ultimately, I will argue, the value of self-trust shows signs of encroaching on the strictly epistemic question of when our beliefs may be said to be justified. Hence again, even if proponents of impartiality are right about what is rational, they may be wrong about what knowledge requires.

First I will clarify some of my terminology and assumptions (Sect. 2), then I will explain the sort of value I take self-trust to have (Sect. 3), then I will give two arguments for the main thesis (Sects. 4, 6) and consider some objections (Sects. 5, 7, 8, 9, 10). Finally I will consider the possibility that self-trust encroaches on epistemic justification (Sect. 11).

## 2 Terminology and assumptions

Fundamental to my discussion is the notion of *rationality*. What counts as rational in a broad sense depends on one's goals. In a narrower sense, however, it is customary to speak of the rational as what one ought to believe in light of one's evidence, given the aim of believing what is true and not believing what is false. This is what I will refer to as *epistemic rationality*.<sup>1</sup> One might naturally assume that considerations of epistemic rationality determine both what we ought to believe and what qualifies as epistemically justified. I mean to call into question both of these assumptions.

---

<sup>1</sup> There is no canonical definition of rationality, or even of epistemic rationality. The core idea here comes from Foley (1987). For a helpful discussion of the varieties of rationality, see Kelly (2003). On Kelly's account, one should be called epistemically rational "when one believes propositions that are strongly supported by one's evidence and refrains from believing propositions that are improbable given one's evidence" (2003, p. 612). The arguments of this paper could all be adapted to that conception of rationality.

Disagreements come in many kinds, so let us here have a sample case, to help focus our intuitions.

Mike and Elinor disagree about politics. Mike thinks that the role of government should be kept to an absolute minimum, and that in particular it is wrong for government to tax the wealthy for the sake of helping the poor. He thinks it violates the rights of individuals for their money to be taken from them in this way. Elinor thinks, to the contrary, that this is among the very most important things that a government ought to do, and that a democratic government violates no individual rights in so doing. Each, with good reason, regards the other as more-or-less equally well-informed and equally intelligent. Indeed, each has talked extensively with the other, and is familiar with the other's reasons, but still is utterly convinced that the other is wrong.

I will refer to cases of this kind as *peer disagreements*. The most salient features of such cases are that there are two parties to the debate who each believe (or have a high credence in) incompatible propositions, and that each party has good reason to treat the other as an *epistemic peer*—that is, as being just as well-informed and just as reliable as oneself in cases of this general sort.<sup>2</sup> In any real case, there are a great many other factors that may be relevant. It matters here, for instance, that Mike and Elinor know themselves to be advocating the opposite sides of a long-running, much-discussed, and unsettled debate. But because nothing will turn on the details of this particular case, I will for now rest content with this rough schema.

The most discussed position in the literature on disagreement is generally known as the *equal-weight view*. It holds, very roughly, that in a case of disagreement between you and an opponent, if you take your opponent to be your epistemic peer, then you should give your opponent's view the same weight as your own, and hence (in the simplest case) you should suspend belief on the matter in question.<sup>3</sup> (Generalizing from the special case of equal weight, I will sometimes speak of *epistemic impartiality*, by which I mean giving equal consideration to the views of oneself and others.) In the sample case, then, both Mike and Elinor should abandon their deeply held political beliefs and suspend judgment. It has proven tremendously difficult to work out the details of this general approach in anything beyond the most idealized circumstances, and some have questioned whether the view even applies in real-

<sup>2</sup> This formulation is intended to reflect White (2009)'s important observation that 'epistemic peer' should be defined in terms of *expected* reliability. Much of the worries in the literature about the notion of an epistemic peer—e.g., Feldman (2009, pp. 300–301), Frances (2010, pp. 424–425), King (2012)—arise because of the mistaken assumption that an epistemic peer will be someone who is *actually* on a par with regard to information, cognitive abilities, etc. To be sure, one is unlikely ever to meet an actual epistemic peer, and even less likely to know it if one does. But this is not to the point: all that matters is that one has good reason to expect one's interlocutor to be neither more nor less likely to get things right, in the domain at issue. Such a state of affairs is quite common.

<sup>3</sup> Prominent defenses of the equal-weight view include Feldman (2006), Christensen (2007), Elga (2007), and Cohen (2013). Some of its defenders, most prominently Elga (who coined the label 'equal weight') have argued that the view can be defanged of its most intolerable consequences, by showing it to be applicable to a much narrower range of cases than one might suppose. My own proposal amounts to a different path toward avoiding the intolerable consequences.

world, non-ideal cases. Without taking a stand on any of the details of *how* the equal-weight view should be developed, I am going to assume that there is this much right about that impartial perspective: that there are *some* commonly realized situations, having something at least close to the structure of the sample case above, where epistemic rationality does prescribe a strikingly revisionary approach to our doxastic practices, in that it prescribes the suspension of belief.

I am formulating this assumption in as general terms as I can, so as to make it as palatable as possible. Of course I need more than the trivial claim that, sometimes, learning about disagreement makes it rational to suspend belief. Everyone can accept that. I want the stronger claim that there are cases of this sort along the lines of the sample case, where the equal-weight verdict is seriously disruptive to how we ordinarily go about forming beliefs. To say that there are *some* such cases, without attempting to specify what they are, skirts a great many contentious details. The assumption is still controversial, inasmuch as some have argued that the equal-weight view is categorically wrong about cases of peer disagreement. But at least many parties to the debate either accept something in the neighborhood of the equal-weight view, or at least concede that this view is right about some cases.<sup>4</sup> Moreover, readers who put themselves in the categorical-denial camp should still find the assumption worth taking seriously, at least for the sake of argument, because I wish to use the assumption to motivate a way of avoiding the unpalatable consequences that seem on their face to arise from impartiality. If I am right about how those consequences can be avoided, then the categorical-deniers, insofar as it was the awful consequences that motivated them, will have reason to reconsider their opposition.

### 3 Self-trust and other doxastic values

It is uncontroversial that, in some sense, we value self-trust. This is to say that we think it a good thing for people to have confidence in themselves in various domains—that they trust the way things seem intellectually and perceptually, as well as morally and emotionally. When epistemologists have considered such matters, they have generally done so from the perspective of epistemic rationality: that is, the question of why self-trust is good has been understood as roughly the question of how self-trust contributes to a greater proportion of true over false beliefs.<sup>5</sup> And there is no doubt that self-trust can have this sort of instrumental value. But the hypothesis I wish to consider is that self-trust has intrinsic value as well.

To get clear about what it would mean for self-trust in the cognitive domain to have intrinsic value, consider Maureen: intelligent, well-informed, confident. There

<sup>4</sup> A prominent example of this sort of mixed verdict is Lackey (2010).

<sup>5</sup> Here, and throughout the paper, this sort of formulation is only roughly right, because it is often suggested that, even from a narrow epistemic point of view, there is more that matters than just getting at the truth (and avoiding error). For discussion see, e.g., the papers collected in Steup (2001). For more narrowly epistemological discussions of self-trust, see e.g. Lehrer (1997) and Foley (2001).

is value in how many true beliefs she holds, and how few false beliefs, and how those beliefs are aptly proportioned to her evidence. But in addition to her goodness as an epistemically rational agent, there is the further goodness of her confidence, something that—according to my hypothesis—is of independent, intrinsic value. Of course it would be natural to suppose that this confidence is valuable only insofar as it serves to further her epistemic rationality (a possibility to which I will return later). But suppose that Maureen's confident self-trust in fact served to further a worldview that was pervasively false and deeply irrational. Still, on the hypothesis I am considering, the self-confidence itself would have some kind of intrinsic value. At least that part of Maureen's cognitive make-up ought to be judged valuable.

Clearly there are many different ways in which self-trust might manifest itself, even if we consider only the cognitive domain. For present purposes, however, we can restrict our attention to one fairly well-defined instance of self-trust, the case where an agent puts greater weight on the way things seem to her in the face of peer disagreement. We can say of such a case that her action exhibits self-trust; we may even suppose her to have an underlying disposition toward self-trust. In Sect. 6, I will offer a more precise account of what such praiseworthy self-trust consists in.

The self-trust hypothesis can be seen as part of a broader hypothesis asserting the existence of doxastic values that extend beyond the concerns of epistemic rationality. Some such broader values are familiar enough. In particular, epistemologists routinely acknowledge the existence of pragmatic reasons for holding beliefs, as when the belief that you will recover from illness helps in your recovery.<sup>6</sup> Other sorts of familiar values are less obviously instrumental. As is often remarked, we value not just any true belief, however trivial, but true beliefs that are judged as in some way *important*. Might there be many such values, beyond the narrow dictates of epistemic rationality, that influence our views about when it is proper to hold a belief? The broader hypothesis holds that this is indeed the case. Self-trust, on this picture, is just one of a family of commendable doxastic dispositions, such that beliefs formed on their basis are to that extent more valuable than they would otherwise be. What other such doxastic values might there be? One might think of trusting one's friends and family, trusting God, and trusting certain institutions.<sup>7</sup> Generalizing from these examples, one might suppose that trust itself is of doxastic value. One might also value creativity, or orthodoxy, or unorthodoxy.

---

<sup>6</sup> Typically, these are noted only to be set aside as not relevant to epistemology. See, e.g., Sosa (2007, p. 44): "First, we are interested in distinctively epistemic justification, not in pragmatic or any other sort of justification of beliefs or other attitudes."

<sup>7</sup> Although I am unaware of anyone who has argued for self-trust as having non-epistemic doxastic value, some excellent work has been done on the subject of trust in others. On the value of putting trust in friends, see Keller (2004) and Stroud (2006), who concludes that "friendship requires epistemic irrationality" (p. 518). Adams (1987) has made somewhat similar claims in a religious context, and Baier (1980) has argued for a secular faith that puts trust in both individuals and institutions: "the secular equivalent of faith in God, which we need in morality as well as in science or knowledge acquisition, is faith in the human community and its evolving procedures—in the prospects for many-handed cognitive ambitions and moral hopes" (p. 133). I discuss the doxastic value of hope in Pasnau (forthcoming-b). For a recent attempt to account for this sort of trust in the context of disagreement, see Hazlett (2014), though his proposal is quite different from my own, because he attempts to defend the epistemic rationality of such practices.

One might think it commendable to remain steadfast in one's views, or commendable to be flexible. To treat these as candidates for being of doxastic value is to propose that they have a value that is not merely pragmatic, nor limited to their propensity to get at the truth. Their value is intrinsic. In what follows I will not make any assumptions about the best way to conceive of such values: whether, for example, they should be thought of as epistemic virtues, or should be captured in terms of rules or permissions. I will, moreover, confine myself to the case of self-trust in the narrow context of peer disagreement. Let the reader be put on notice, however, that this one case is just the thin wedge of a larger agenda.

Of course, all these values may be outweighed by other values. Trust, for instance, can be misplaced, or taken too far. To say that it is of intrinsic value implies that it is good, wherever it is found. But that goodness may easily be outweighed by other factors, as when trust is so obviously irrational as to be, on balance, quite blameworthy. Notice, however, that the same is true of epistemic value. If I find my son spending all afternoon at the computer, mastering the nuances of *Minecraft*, I will not feel pleased at all he has learned, even if he has acquired more true beliefs that way than he might have by doing his Spanish homework. An account of how exactly to weigh these countervailing values is beyond me. But I will later (Sect. 8) consider the question of why self-trust seems more praiseworthy in some contexts than in others.

Right away, my proposal may look like an obvious non-starter inasmuch as it appeals to what sometimes gets called the “wrong kind of reason” for belief: a reason that does not make the belief more *rational*, or more likely to be *true*; a reason that focuses not on the *object* of the belief but on the desirability of the *belief-state*; a reason that is in fact not a true *reason* at all to have a particular belief, but at most a reason to *want* to have such a belief. It may even look, right from the start, as if it would be impossible—or at any rate blatantly contradictory—to hold explicitly the conjunction of views I recommend. This last objection is one that I will have to confront head on (in Sect. 7), but I am not in general bothered by the charge of appealing to the wrong sort of reason. It is in fact precisely my ambition to argue that reasons of this kind—not truth-conducive, not grounded in features of the object of belief—can sometimes count as good, intrinsic (not merely pragmatic) reasons for belief.<sup>8</sup>

<sup>8</sup> The main text here canvasses very quickly some of the tests that have been advanced to distinguish between the right and the wrong kinds of reasons. See, e.g., amidst a large literature, Schroeder (2012b) and D’Arms and Jacobson (2014). The latter makes it clear just how murky this distinction is, and I scarcely wish to put much weight on it here. But it is perhaps worth noting that, on its face, self-trust passes at least one prominent test for being the *right* kind of reason. For it scores quite high as being the sort of reason that people are able to respond to as a reason for belief. Threaten someone to believe *p* or else, and the threat likely will not work. (That’s the wrong kind of reason.) In contrast, tell someone who is tempted by the arguments for the equal-weight view that she should have the courage to trust herself, and this is as likely as anything to have some influence. It looks, in this respect, like the right kind of reason. But this may show simply how unreliable that particular test is, since one might suggest that various reasons that are clearly of the wrong kind, such as various forms of bigotry, will fair quite well on the “ability to respond to” test. Alternatively, one might try to save the test by pointing out that self-trust motivates belief only indirectly, by leading an agent to put more weight on her own first-order reasons, reasons that are presumably of the right kind. (This will become important in Sect. 7.)

Lest this seem right from the start to be giving away the game, let me here give a quick argument for why such a thesis in general should not immediately be dismissed as indefensible. For suppose, to the contrary, that someone thought the only good (non-instrumental) reasons for belief with regard to a given question were the usual sorts of object-given pieces of evidence. Line these pieces up on each side of the question, pro and con, giving each reason whatever weight it seems to deserve. If these were the only (non-instrumental) factors that determine belief, then what we ought to believe should simply fall out of this array without further ado. But plainly it does not. Should we simply place our belief in whichever side has greater weight? What *reason* tells us to do that? Presumably, however, we would not do that, but would suspend belief unless the reasons on one side outweighed the other side by a sufficient margin. But how much greater weight must one side have? What *reason* is there for choosing a certain threshold rather than another? Of course, the higher the threshold, the greater the likelihood of truth, and in that sense such considerations may be said to be truth-conducive. But no one (except perhaps the skeptic) thinks the threshold should be set at perfect certainty, nor does anyone think we should believe everything that is more likely than not. Evidently, there is more at stake here than simply the epistemic goal of obtaining a greater proportion of true over false beliefs. Whatever those further values are, my point is only that we do in this area rely on values that are not in any straightforward way the right kinds of reasons. Now, some might respond to such considerations by denying that belief should be treated in binary, yes-or-no terms, and arguing instead that we should think of belief as coming in degrees (or credences). On this approach, no question of threshold arises. But it is a familiar phenomenon that we do, in everyday life, fully commit ourselves to the truth of various propositions, in the sense that we accept them without doubt, even in cases where the evidence is merely good enough and not strictly indubitable. For such perfectly ordinary practices to be legitimate, there needs to be room for reasons of a kind that go beyond the evidence, explaining why it is legitimate to believe on the basis of *this much* evidence, but to suspend belief where only *that much* is to be had.<sup>9</sup>

#### 4 The initial argument

Now that the stakes are clear, I had best hasten to make my case. The hypothesis in question makes a claim about value: that beliefs based on self-trust have intrinsic value, just in virtue of being so-based. This is to say that beliefs so-formed are praiseworthy, at least insofar as they are so-formed, and that we have reason to form beliefs in this way. If these normative claims are right, then we have a potential solution to the impasse over disagreement. We can say that in a case of peer disagreement, it would be epistemically rational for each party to give the other

---

<sup>9</sup> The argument of this paragraph is indebted to Schroeder (2012a, p. 276), who argues somewhat similarly that “reasons to withhold can’t be evidence.” But the reasons he is interested in are pragmatic ones, not the sorts of intrinsic values I am concerned with, and his argument arises in a wholly different context. For more on credences versus full belief, see note 22.

side's position equal weight, and so suspend belief on their political views. But we can also say that this result appears intolerable because there are other things that are valuable, where belief is concerned, beyond epistemic rationality. So, returning to the sample case, even while Mike's rationality would be admirable, if he gave Elinor's views equal weight, his lack of confidence in his own views would be regrettable, perhaps even blameworthy. Of course, everything rests on the precise details of particular cases, and so sometimes epistemic considerations may loom larger than considerations of self-trust. But in a general way we can now understand the root of the impasse.

So one way to state the argument for my hypothesis is that it explains the impasse over disagreement. But one might fairly object that this hardly counts as an argument at all, inasmuch as I have simply assumed from the start that there is an impasse. I am going to continue to assume that there are cases where the equal-weight view gets things right. But I do not wish simply to assume that the consequences of impartiality are intolerable. So the heart of my argument will be to make the case that the consequences are bad, and to suggest that the unpalatableness arises, at least in part, because of the way it conflicts with the value of self-trust.

As an initial argument, I will try to make a persuasive case simply that we—you and I, and people in general, and philosophers in particular—do in fact care about self-trust, and that we do so independently of other things that we care about. Whether such an argument yields a normative conclusion is an open question, but I will reserve that issue for the following section.

Let me start with myself. I think it is quite unreasonable to demand that either Mike or Elinor suspend their belief on political matters. I myself hold fairly strong political views, of familiar sorts and with the usual sort of intensity. I might easily put myself into this story of peer disagreement, but even so, I am not about to give up *my* political views—it is not even a serious possibility for me. So when the equal-weighter suggests that I must do so, I feel very confident that something is being left out. *Maybe* such impartiality gets the story wrong as regards epistemic rationality. But if I continue to suppose that, at least in *some* such stories, the equal-weight view is right about that, then I see no other conclusion other than that there is more that matters in forming beliefs—at least to me—than epistemic rationality.

Now let me work on you, and find out what sorts of controversial matters you care deeply about. Does reflection on the phenomenon of peer disagreement make you feel as if you should give up those beliefs? There are smart people on all sides of these questions. Perhaps you hold the majority view, or perhaps a minority view. Do such considerations of number matter? Do you think that you should adjust your credences in proportion to the views of others? Does it not feel quite wrong to suppose you must give your beliefs over to the consensus view, or suspend belief, ignoring how in fact things seem to you?

I am aware that not everyone will be moved by these rhetorical questions. Committed equal-weighters have made it quite plain that they are prepared to follow the argument where they take it to lead, and that they are unpersuaded that any intolerable consequences loom. David Christensen, for instance, regards impartiality as “pretty good news,” inasmuch as it yields a “welcomed” and “valuable strategy



for coping with our known infirmities.”<sup>10</sup> Perhaps this is your view too, and if so then I am unlikely to persuade you otherwise. But if this is your view then I think you, like Christensen, will have to work hard to maintain it against the very stubborn intuition that there is something problematic about simply throwing over one’s beliefs in these domains. Such intuitions are very widely held, even among professional philosophers. Consider Philip Pettit, who remarks the consequences of giving equal weight “would be objectionably self-abasing” and amount to “servility.”<sup>11</sup> Or Adam Elga, who charges that it would lead to “spinelessness” and a “lack of self-trust.”<sup>12</sup> Or Ernest Sosa: “it would be bad to have to suspend judgement on just about any controversial question.”<sup>13</sup> None of these authors stops to consider the basis for their normative claims. What they all officially *argue* is that epistemic rationality does not require impartiality in any widespread way. None of them shows any signs of allowing other sources of doxastic value that might have a bearing on the case. And yet what they actually *say* seems to invoke self-trust as an independent doxastic norm. This is the intuition that seems to motivate each of these authors to look harder at the epistemic calculus, and find some basis for escaping the consequences of giving equal weight. My hypothesis is simply that we take this intuition at face value and recognize it for what it is.

In one vital respect, the passages just quoted make a stronger claim than I seek to establish. Call the *strong hypothesis of self-trust* the thesis that it would be positively wrong, in at least some cases of peer disagreement, to be so “self-abasing” (Pettit) as to suspend belief. Maybe so, in some cases. But all I seek to establish is the *weak hypothesis of self-trust*, that the doxastic value of self-trust has sufficient weight to make it the case that it would not be wrong, at least in some cases of peer disagreement, to give significantly less weight to the views of one’s epistemic peer. The strong hypothesis, then, tells us that self-trust *requires* certain credences, whereas the weak hypothesis maintains only that it *licenses* certain credences.

The strong hypothesis, although tempting, is too strong. We should agree with Christensen that there would be something admirable about someone who, in the face of peer disagreement, decided to suspend judgment on all the controversial matters she had hitherto committed herself to. We should not blame such a person for acting in this way, nor should we call it “servile” (Pettit) or “spineless” (Elga). We should not suppose that the value of self-trust is so great as to overwhelm considerations of epistemic rationality entirely, and demand that one hold steadfast in one’s views.<sup>14</sup> But at the same time we should reject the equal-weighter’s

<sup>10</sup> Christensen (2007, p. 216). For a similar verdict, see Kornblith (2010, pp. 51–52). Even Kelly, a leading opponent of the equal-weight view, remarks that he does not regard the implication that we should give up our beliefs in these controversial domains as “a significant theoretical cost”; on the contrary, the suggestion we are too confident in these domains has “considerable independent plausibility” (2010, p. 128).

<sup>11</sup> Pettit (2006, p. 181).

<sup>12</sup> Elga (2007, p. 484).

<sup>13</sup> Sosa (2010, p. 284).

<sup>14</sup> It should be noted that the motive of self-trust hardly could yield such a strong claim, inasmuch as agents who are persuaded of the equal-weight case should presumably trust *that* conclusion, and should

contrary claim that it would be wrong to maintain one's original beliefs in the face of disagreement. Even if such behavior is epistemically irrational—as I am assuming—it does not follow that holding steadfast would be wrong. Even on this weaker version of the hypothesis, there turn out to be doxastic values that obtain independently of epistemic considerations.

Is self-trust the critical doxastic value at work in such cases? Both Pettit and Elga speak as if it is, and this is what my hypothesis of course maintains. But someone might well see a more complex set of doxastic values at work. Perhaps Mike's steadfastness in the face of Elinor's arguments is motivated not just out of self-trust, but out of a sense of solidarity with his family, or his faith community, or an old teacher, or even out of a sense of solidarity with his past self, which for years, we might imagine, has been championing conservative political views. Insofar as you suppose that any such factors are at work, you are embracing the broader hypothesis that I described in the previous section. In Sect. 6, I will consider a reason for thinking that *self-trust* has special value in the context of disagreement. But I am not adverse to casting the paper's main argument in a broader form.

## 5 Is versus ought

Before considering the further argument for my hypothesis, we should consider an obvious objection: that the argument just made simply reports on how *in fact* some people conduct themselves doxastically, and shows nothing about how people *ought* to regulate their beliefs. It is precisely because I take this objection so seriously that I have been describing my thesis only as a *hypothesis*. It is entirely possible, for all I will be able to show, that I am capturing nothing more than one aspect of our deplorable tendency toward irrationality in the way in which we form beliefs, a tendency so ingrained in us that it strikes us a good thing, but a tendency that—especially as philosophers—we should fight against with all our intellectual strength. In honor of this possibility, let me draw a further distinction among positions. Let the *descriptive hypothesis of self-trust* be that our general pre-theoretical failure to give equal weight to our epistemic peers, and even the widespread philosophical resistance to such impartiality, in fact arises out of a tacit commitment to the value of self-trust. Let the *normative hypothesis of self-trust* be that such an attitude of self-trust, even in the face of peer disagreement, is in fact a good thing. The thesis of this paper, then, strictly, is the weak normative hypothesis of self-trust.

Even the descriptive hypothesis is of considerable philosophical interest, because it offers to explain, at least in part, why the phenomenon of disagreement has caused such perplexity. Although the general idea of epistemic impartiality is scarcely a difficult notion to grasp, almost no one is prepared to follow this advice. In a way, this is the most intriguing issue in the whole debate: why does almost no one give equal

---

Footnote 14 continued

follow their inclination to suspend belief. But such agents might themselves be threatened with epistemic irrationality, inasmuch as the equal-weight view is itself the sort of disputed thesis where suspension of belief seems mandated. On the self-referential character of the equal-weight view, see Elga (2010).

weight, despite the obviousness of such a course of action? The descriptive hypothesis offers an explanation. All of us, ordinary folk and philosophers, feel the pull of doxastic values that restrain us from achieving perfect epistemic rationality. Philosophers generally feel the wrongness of the equal-weight result, and so cast about for some failure in the argument. Ignoring the existence of non-epistemic doxastic values, they look for one or another subtle way in which impartiality goes wrong in its calculation of epistemic rationality. The descriptive hypothesis suggests that these discussions are a mere distraction, and that in fact the real issue is the abiding sense that there would be something “bad” (Sosa) about giving equal weight. If equal-weighters wish to persevere in this debate, they need to focus not so much on the narrow epistemic issues, but rather on broader questions about the ethics of belief.

What about the prospects of moving from this descriptive hypothesis to the normative hypothesis that we *ought* to give some non-epistemic value to self-trust? Here everything rests on how seriously one is willing to take the intuitions I described in the previous section. That in turn depends not just on whether one shares those intuitions, but also on whether one is prepared to entertain the broader hypothesis that there are non-epistemic doxastic values. There is an extraordinary sort of doxastic purity that prevails over the field of epistemology, according to which the only norms that should govern belief are those of epistemic rationality. But behavior—even doxastic behavior—can be rational without being epistemically rational, so long as there is some other, non-epistemic goal. As noted earlier, even the most puritanical epistemologist will allow that in certain special cases there may be good pragmatic reasons to form beliefs that, from a strictly epistemic point of view, would be irrational. I also argued (at the end of Sect. 3) that we should take explicit account of the sorts of values that determine how strict or lax we should be in forming beliefs. Such questions of threshold are not governed in any straightforward way by the ambition of maximizing true over false beliefs, but still they seem to depend, at least in part, on values that concern the intrinsic character of belief. Few, for instance, would approve of believing a proposition judged to be only 51 % likely. Doubtless there are pragmatic reasons why this would be a bad idea, but it also just looks like an intrinsically blameworthy doxastic practice. This suggests that epistemology ordinarily extends to various non-epistemic values that are intrinsic to our conception of belief formation.<sup>15</sup> The

---

<sup>15</sup> An anonymous referee suggests that epistemic values alone might account for why the threshold for belief is greater than 51 %, inasmuch as our epistemic values might simply give more weight to avoiding false belief than to achieving true belief. Yet still the question arises of why we weigh one of these twin desiderata more than the other, and to what degree. It is reasonable enough to think that the twin desiderata are simply basic epistemic values, with nothing more to be said about what makes them valuable. But it would be odd to think that the decision to weigh one more than the other is likewise basic. On the contrary, the decision to give more weight to avoiding falsehood seems to call out for some sort of further explanation, grounded in some sort of further value. Of course there will be many *pragmatic* reasons, in differing circumstances, for weighing the risk of error more or less heavily. But here the question is what sort of *intrinsic* weight is to be placed on one or the other desideratum. It is perhaps merely a terminological point whether such values get included among our “epistemic” reasons, but for my purposes what matters is just that we need to expand the field of doxastic values beyond what is usually counted as epistemic.

further step I wish to take countenances such reasons even when they *conflict* with our narrowly epistemic values. Such conflict threatens to engulf my proposal in contradiction. But before turning to that objection, let me first introduce the promised further argument for my hypothesis.

## 6 The further argument

So far I have argued only that we should endorse the weak hypothesis of self-trust because it captures a widely-shared intuition. The case would be more persuasive, however, if I could better articulate exactly why self-trust is supposed to have intrinsic value in cases of peer disagreement. Let me try to do so.

We can distinguish two ways in which you might come to give equal weight to the opposing view of an epistemic peer. In one way, you might listen to your opponent's reasons, and come to feel the force of those reasons. In that case, supposing your original reasons continue to have force, it would seem strange for you not to give your opponent's reasons equal weight, and so it would seem right for you to suspend belief on the matter in question. (As usual, I assume the simplest and most idealized sort of case.) It would be strange to do anything else, because we are imagining that you yourself have come to feel the force of your opponent's arguments.

In another kind of case, your opponent's reasons leave you entirely cold. In our sample story, for instance, Elinor might be quite unable to feel the force of Mike's appeal to the property rights of the individual against the state. To be impartial here, Elinor has to think to herself along the lines familiar from the debate over peer disagreement: that Mike is likely to be just as smart as her, every bit as well informed, and so forth. With those thoughts foremost in mind, she gives equal weight, even though the reasons for the other side have no pull on her at all.

It is this second kind of case that seems problematic, and it is here that self-trust reveals its independent, non-epistemic value. In a case such as this, the agent must give significant credence to a thesis whose merits she cannot herself directly recognize. From a certain perspective, this is a smart thing to do, because, as the equal-weighters insist (and as I continue to assume), there are strong reasons to think the perspectives of others are as likely to be right as one's own. If all one cared about were truth, and one were simply playing the odds, then one ought to give credence to their reasons, even if they leave one completely cold. From a broader perspective, however, there seems something undesirable about such a course of action. What we value, in the cognitive domain, is not just forming the right beliefs, but forming them for the right reason, with an understanding of why they are true. In the first kind of case above, an agent would be doing just this, suspending belief because he understands the equal and opposed strength of the arguments on each side of the question. Because he himself feels the force of the considerations on each side, suspending belief is a way of trusting his own judgment. In the second case, in contrast, the agent must reach a conclusion that is not borne out by her own judgment about the two cases. As she sees things, the first-order evidence is decisive on one side, and yet giving equal weight requires

that she go against the way things seem to her, and reach a conclusion without any real understanding of why each side is intrinsically plausible. Even if suspending belief is the best bet, from a purely epistemic point of view, still it is a conclusion that requires her to divorce her credences from the way things actually seem to her. Here lies the precise sense in which failing to have self-trust seems objectionable in the context of peer disagreement.<sup>16</sup>

This line of thought helps explain why the proper conclusion to draw is not the strong hypothesis but only the weak one (see Sect. 4). It is not that Elinor, in the sample case, would be doing something wrong by suspending belief in the face of Mike's arguments. For even though she cannot feel the force of his first-order arguments, she can feel the force of the equal-weighters' second-order arguments, and it may be that she will "welcome" this sort of outcome as "pretty good news" (Christensen). She would then, in one respect, be trusting her own judgment about how to reason in such cases, and forming her credences on the basis of those good higher-order reasons. When the story is so told, it is hard to see what grounds for objection there could be. But we can also see, in the present context, what licenses Elinor to stand fast in her views, despite the equal-weight argument. Giving equal weight to Mike requires giving significant credence to a view to which she cannot herself feel any attraction. Her only evidence for it is the higher-order evidence that comes from Mike's testimony about how things seem to him, conjoined with the rationale of the equal-weight view. Wholly lacking in any first-order understanding of what makes Mike's position compelling, epistemic impartiality would require her to alienate herself from how things seem to be. There is something clearly unsatisfactory about such alienation, and its badness is not merely pragmatic, but intrinsic to what Elinor is being asked to do. We ought not to blame someone for refusing to do it.

## 7 Believing what is less likely to be true

The further argument just described brings the value of self-trust closer to the sorts of concerns that epistemologists have always recognized, inasmuch as it criticizes the equal-weight view for divorcing one's credences from one's evidence. But it does so in a way that is inconsistent with epistemic rationality, inasmuch as it allows

---

<sup>16</sup> I am indebted here to the work on moral deference in Hills (2009) and McGrath (2011). Compare McGrath's remark, regarding someone who defers to a moral expert, that "even granting that such an agent might very well deserve moral praise, he or she still falls short of an important ideal associated with moral agency: that of doing the right thing for the reasons that make it right" (p. 135). Both Hills and McGrath explicitly contrast the moral and cognitive domains, and apply their conclusions only to the moral. Here, then, I am suggesting a way in which the general structure of their argument can be extended even into the cognitive domain. Compare Sliwa (2012), who likewise denies that moral deference is a special case, and then uses that claim to argue for moral deference's legitimacy. I take the considerations here to point toward a sharper understanding of the contexts in which deference (moral and otherwise) is problematic.

first-order seemings to take precedence over the second-order argument for giving equal weight. Self-trust in such contexts thus violates one of the chief principles of epistemic rationality, that what we ought to believe should reflect our total evidence about what is most likely true. One familiar version of this general line of thought appears in John Locke: “It is very easily said, and nobody questions it, that giving and withholding our assent, and the degrees of it, should be regulated by the evidence which things carry with them.”<sup>17</sup> In Bertrand Russell, this idea often gets expressed with a nearly religious fervor: “Morally, a philosopher who uses his professional competence for anything except a disinterested search for truth is guilty of a kind of treachery.”<sup>18</sup> Although Locke is focused on evidence, and Russell on truth, I take it that these claims amount to much the same thing: that we ought to regulate our beliefs strictly in accord with how our evidence suggests that things are likely to be.

Let us use the term ‘evidentialism’ for the principle that one’s beliefs should always be proportioned to one’s evidence. Plausible though it may be, this principle is inconsistent with my hypothesis, which holds that even when one recognizes that the total evidence favors suspending belief, it can still be acceptable to privilege one’s own beliefs, letting those beliefs be “regulated” (Locke) not by all the evidence but by a subset that has some special connection to oneself. This way of pursuing truth is by no means “disinterested” (Russell), but rather privileges one’s own point of view in a way that the total evidence does not support. To be sure, there are many distinguished philosophers who have questioned evidentialism, in contexts as various as religion, morality, and rationality itself.<sup>19</sup> My hypothesis entails that such anti-evidentialism also applies in cases of peer disagreement.

It is of course impossible here to do justice to the very large literature that surrounds the thesis of evidentialism. But I do need to say something about one particular kind of argument that is commonly made in its favor, because this objection might seem particularly relevant to the context of peer disagreement. Jonathan Adler, among others, has argued not just that evidentialism is true as a normative constraint on belief, but that the normative thesis is trivially true, inasmuch as it describes a constraint that it is not even *possible* for us to violate. More cautiously, Adler thinks that we cannot self-consciously be in a position where we both believe a proposition and believe that one’s evidence is out of line with that belief.<sup>20</sup> This seems plausible, on its face, and it looks to have direct application to

<sup>17</sup> *Conduct of the Understanding* Sect. 33. See also *Essay IV.17.24*.

<sup>18</sup> Russell (1945), p. 835.

<sup>19</sup> See note 7 for various cases. And see e.g. Foley (1993) for the rejection of evidentialism in favor of a broader conception of rationality. ‘Evidentialism’ is used in a variety of different ways, but here I roughly follow, e.g., the usage in Chignell (2010): “the position, roughly, that we are obliged to form beliefs always and only on the basis of sufficient evidence that is in our possession” (Sect. 1.1). On other characterizations, evidentialism is framed in terms of the *epistemic rationality* of conforming one’s beliefs to the evidence, and when so framed I of course have no quarrel with the doctrine.

<sup>20</sup> See, e.g., Adler (2002, p. 25): “Any account of the ethics of belief should fit tightly with the crucial fact that it is not possible to regard oneself as both holding a belief and holding that one’s reasons for it are inadequate.” For another argument that reaches a similar conclusion, see Shah (2006), who holds that “truth is hegemonic with respect to doxastic deliberation” (p. 490).

the self-trust hypothesis. Consider that, in a case where you and I disagree over  $p$  in some domain rife with disagreement, and where I regard you as my epistemic peer, it seems that I would accept the following propositions all at the same time:

1. I believe that the likelihood of  $p$ 's being true is 0.5 (given that we are locked in peer disagreement, and given that I accept the equal-weight view as a matter of epistemic rationality)
2. I nevertheless believe  $p$  (my antecedent view, which I maintain on the grounds of self-trust)
3. I believe that  $p$  is true (forced by 2)
4. I believe that  $p$  is very likely to be true (forced by 3)

But of course (4) and (1) are inconsistent. I cannot abandon (4) without abandoning the self-trust that led me to (2). I cannot abandon (1) without dropping my belief that the equal-weight view gets things right as regards epistemic rationality. Hence it may seem I cannot consistently embrace the equal-weight view as a guide to epistemic rationality while maintaining my own view out of self-trust.

What I must do, to avoid contradiction, is reject (1). I do not believe that the likelihood of  $p$ 's being true is 0.5—I believe instead, as (4) has it, that  $p$  is very likely to be true. But then in what sense do I accept any part of the equal-weight view? Well, what I accept is that—if all I cared about were tracking the truth on the basis of all the available evidence—then I should suspend belief in  $p$ , regarding its likelihood as 0.5. But since this is not all I care about, the consequent does not follow. Other considerations motivate my beliefs—in particular, I give pride of place to how things seem to me at the first-order level, trusting myself more than others. Hence I refuse to believe that which, if I cared nothing about self-trust, I would believe.

In thinking along these lines, self-trust does not directly furnish me with a reason to believe  $p$ . It is not as if I think:  *$p$  must be true, because I believe it!* Rather, my tendency to trust myself leads me to give special weight to the reasons that strike me as correct, and to downplay the force that rationality would require giving to the reasons of others. My first-order reasons are what motivate me to believe. In giving this special weight to one's own first-order reasons, the account approaches a prominent line of argument made against equal-weight views: that such views ignore the evidence.<sup>21</sup> But the difference is that my account makes no claim for the rationality of focusing specially on one's own reasons. On the contrary, it concedes to the equal-weight view that, from a strictly epistemic point of view, the reasons of others should be treated with impartiality. And it asserts that even an agent who recognizes that this is so both can and ought to give special weight to her own reasons.

But *can* someone really think this way? I believe that we do it all the time. We regularly form beliefs even while recognizing that, from a certain point of view, the odds of being right do not look good. But is that not to say that we believe  $p$  even

<sup>21</sup> See e.g. Kelly (2010) and Enoch (2010): “the Equal Weight View requires that in the face of peer disagreement we ignore our first-stage evidence altogether” (p. 969). For criticism of the neglected-evidence objection see Pasnau (forthcoming-a).

while believing that  $p$  is not likely to be true? No—that would be contradictory. But it is not contradictory to believe a thing even while recognizing another point of view—one grounded, perhaps, in the cold, hard evidence—from which that thing would seem unlikely. Philosophers in particular, I believe, do this sort of thing all the time, when they adhere to what seems right to them, knowing full well that many—maybe most—will disagree with them. Perhaps you will suspect that I am doing just this, right now. *Nolo contendere*. After all, I myself accept that the equal-weight view gets things right as a matter of epistemic rationality. Moreover, I am aware that my self-trust hypothesis is controversial. Still, I believe this hypothesis all the same.

So much the worse for me, you may conclude. But now let me suggest another context in which we regularly do something similar: when we fully believe things for which we have highly probable but non-certain evidence. Full belief, as I use the term, is belief without doubt, and at the end of Sect. 3 I argued that our rules for deciding on the threshold of where belief should begin are an instance of the sort of intrinsic, non-epistemic doxastic value I am arguing for. If that is right, then it should be no surprise that such values can conflict with our narrowly epistemic norms, and so it is. Imagine your evidence puts the likelihood of  $p$  at 0.99—that is, almost certainly true. Suppose, as a result, that you fully believe  $p$ . Such a state of affairs seems quite common. John Locke described how some propositions “border so near upon certainty that we make no doubt at all about them, but assent to them as firmly ... as if they were infallibly demonstrated” (*Essay* IV.15.2). As Locke is well aware, this violates the evidentialist rule. Indeed, it threatens to lead to a contradiction of just the sort considered above in the case of self-trust. Now there are many extant proposals for either doing away with full belief or understanding it in a way that does not threaten epistemic rationality. But here my point is only that—whatever may be theoretically optimal—in fact people often do form full beliefs in a way that requires them to set aside a perspective that better conforms to the evidence. It is a similar sort of doubleness in attitude that one finds in the case of self-trust.<sup>22</sup>

## 8 Why is self-trust not always praiseworthy?

If self-trust has intrinsic value, as my hypothesis maintains, then it seems as if we should always value it. But in some cases it seems that we value it not at all. If I

<sup>22</sup> Arguing that the case of full belief is analogous to self-trust invites the thought that perhaps my theory might best be advanced by *directly* appealing to the distinction between credence and full belief. Whether this will help depends on one’s theory of belief. On views that divorce credence and belief, it is relatively easy to invoke doxastic norms that depart from epistemic rationality. Indeed, Kaplan (1996) goes so far as to contend that “there is no reason why we cannot rationally believe the improbable” (p. 148). In contrast, on views that identify belief with sufficiently high credence (e.g., Hawthorne and Bovens 1999), there is no gap between belief and credence to be exploited. Indeed, on such views there is no basis for arguing that full belief violates epistemic rationality, in part because there is no such thing as truly “full” belief, unless an agent’s credence actually hits 1. But this is not a difficulty for the argument here in the main text, which is concerned only with showing that it is psychologically possible for agents to violate evidentialism, and so seeks to show only that in fact the psychology of full belief, in some agents, sometimes involves setting aside evidence that they are aware to be available.



have good reason to think that I am intellectually incapacitated in one way or another—drunk, tired, jet-lagged—then it is proper to discount my conclusions on certain sorts of matters. Self-trust also looks to be wholly inappropriate in certain kinds of disagreement. Here is a case:

David and Kathrin are skiing in the backcountry, where the avalanche risk is high. They recently took a class together on avalanche safety and, as the class taught them to do, they are digging a trench to test the stability of the snow. They observe several very hard layers of snow, sandwiched between softer layers. They disagree, however, on what this means. David seems to recall learning that this is a sign of danger. Kathrin seems to recall learning that this indicates the snow is stable. They compare memories, and find that although neither is completely certain of what they were told, each feels confident of being right. Antecedently, each was inclined to regard the other as an epistemic peer in this domain.

If Kathrin and David were arguing over philosophy or politics, then we would feel the force of self-trust as a doxastic value—or so I have been arguing. But the present case is somehow different. Here, self-trust would clearly be wrong. They should agree to suspend belief and, unless the skiing looks *very* good, they should turn around. (*Caveat lector*: David is right. Layers of soft and hard snow are a sign of avalanche danger.)<sup>23</sup>

Why does self-trust get a grip in the political case, but not in this backcountry case? More generally, why do we sometimes admire self-trust, whereas other times we regard it as stupid, pig-headed, and dogmatic? The obvious explanation, in light of the present case, is that much is at stake on a backcountry skiing expedition, whereas a political disagreement is merely theoretical. But this does not seem to track our intuitions. Even if all Kathrin and David were disagreeing about is how to make cupcakes, we would still put little weight on self-trust. For Kathrin to insist that the recipe said two cups, not one cup of sugar, would seem stupid, not praiseworthy. Conversely, even if our imagined adversaries were debating politics, and were in a position to decide future government policy, I think that many would still give them some credit for holding fast to their beliefs—up to a point, at any rate.<sup>24</sup>

<sup>23</sup> As always, the details of the example matter tremendously. If Kathrin had an extremely vivid memory of learning that a mix of hard and soft layers is stable—if she felt completely certain of it—then she might be so certain she is right that it becomes epistemically rational to doubt David's reliability, no matter how persistently he insists that he too is certain. In such a case we would not need to appeal to self-trust in accounting for why she should carry on despite David's dire warnings. When the story is so told, it becomes an instance of the widely discussed sort of case where one's opponent's view seems flatly absurd, and where equal weight accordingly should not be given. Some regard such cases as a difficulty for the equal-weight view. But I take these simply as cases where it becomes rational to doubt the equal reliability of one's supposed peer. For criticism of the absurd-disagreement objection to the equal-weight view, see Pasnau (forthcoming-a).

<sup>24</sup> See also van Inwagen (1996, p. 139), who considers and rejects the thought that peer disagreement is permissible in philosophy because nothing is at stake. He urges to the contrary that much can be at stake in philosophical disagreement. Someone not ready to give up this low-stakes interpretation of the phenomenon might plausibly suggest, however, that philosophers often do not take seriously enough the real-life consequences their views can have, even when it is true that such views do have important consequences.

A better explanation can be found in the further argument of Sect. 6. If that argument is correct, then self-trust has special value in the context of peer disagreement because it allows agents to connect their credences with their first-order impressions of how things are. The idea was that there is something intrinsically good about forming beliefs in a way that is in harmony with one's evidential perspective. If this is right, then we should expect self-trust to matter more in some cases than in others. If what is at issue is a question you have long reflected on, where you take yourself to have considerable expertise, or whose resolution is central to your self-identity, then the kind of alienation required by epistemic impartiality is a heavy price to pay. Politics matter tremendously to Mike, or so we may suppose, and he has long held his conservative views, and prided himself on his knowledge in the political domain. Should he divorce his credences from how the issues themselves seem to be? Even in full awareness of the epistemic risk he takes, he is surely within his rights to stand by the view of the world that presents itself to him.

Now consider Kathrin, in the backcountry. She feels pretty confident in her memories of the class, but what sort of investment does she have in this? Either she is wrong or David is, and really it would be pointless to insist on her point of view, because nothing in her self-identity turns on correctly remembering this one fact. A great many disputes are like this. In the case of an argument over a recipe, or the location of a restaurant, or the correct way to divide a restaurant bill, it would be nothing more than stupid obstinacy to insist upon one's own perspective. At most there will be a slight chagrin at having to retreat to a view out of line with how things actually seem, but anyone who insisted on his own point of view in such cases would be exhibiting a character flaw, not praiseworthy self-trust.

What about philosophy? If you and I disagree on some hitherto unconsidered question, self-trust would seem dogmatic. We should suspend judgment and reflect on the matter more. For the fledgling undergraduate philosopher, pretty much the whole field should look like this. But as one's thinking matures, it becomes increasingly appropriate to take strong positions. Of course, it is not obligatory: the first-order arguments may not strike you as decisive, or you may cautiously prefer to keep firmly in mind the higher-order story about peer disagreement. But if those first-order arguments do seem decisive, and if you have thought long and hard, and written or taught at length on these subjects, then it would be perfectly appropriate to form certain sorts of beliefs. If one truly cares about the issues, then doing anything else might be a profoundly alienating experience. Standing fast in these circumstances is not dogmatic, but simply a reflection of one's self-identity.<sup>25</sup>

## 9 Might provisional acceptance do just as well?

One way to avoid the self-trust hypothesis, while still accounting somewhat for our intuitions, would be to distinguish between forcefully advocating a position and

<sup>25</sup> For a more wide-ranging discussion of dogmatism, from a rather different point of view, see Kelly (2011).

actually believing it. What is admirable, one might suggest, is simply the advocacy of a view, in cases where such advocacy contributes to a fuller, deeper understanding of the issues. When someone does that we should applaud, but when such advocacy is paired with an actual *belief* in the doctrines being articulated, we should find this not admirable but dogmatic.<sup>26</sup>

It seems to me this fails to do justice to our attitudes. If I learn that some great, idiosyncratic philosopher does not actually believe the views she advocates, I admire her less, not more. If I learn that some intense political dispute at a dinner party was prompted by a guest who did not even believe the things he was claiming, I will be disgusted. No doubt there is a place for advocating views just for the sake of argument. But my impression is that we tend to find this a disappointing facsimile of the real thing. Across a wide range of cases, certainly including philosophy, our admiration increases when we learn that someone is not just setting out a view for the sake of argument, but actually believes it. We admire philosophers who take strong, unorthodox positions—not because we think they are likely to be right, but because we value the boldness and integrity that comes with standing by one's ideas. Now, to be sure, one can display impressive originality without endorsing the products of one's creativity, and such originality is something we admire. But in both philosophy and other intellectual domains, there is something at stake in such cases beyond simply a facility with new ideas. We value people who think for themselves and who have the courage of their convictions, even when, or especially when, those convictions go against the majority. We value not just the provisional acceptance of how things seem, but the forthright belief in those seemings.

To value such self-trust in philosophy—even against the total evidence—does not require a diminished sense of philosophy's importance. It does not turn philosophy into an idle game, nor does it preclude a robust conception of philosophy's capacity to discover the truth. What it does require is a conception of philosophy as having values other than purely epistemic. Part of what makes philosophy valuable is that strong philosophical convictions, powerfully argued for, have a kind of worth all to themselves—they have a beauty and magnificence quite apart from their truth value. We value self-trust in philosophy because we view that as part of what makes for great philosophy, and we value great philosophy whether or not in the end it comes out true.

## 10 Does self-trust have epistemic value?

Could the value of self-trust somehow be grounded in epistemic rationality? One might in particular wonder whether a policy of self-trust advances our long-term epistemic ends, perhaps failing to track the truth in the short run, but in the long run advancing the search for truth. Inquiry is advanced, according to this line of thought,

---

<sup>26</sup> On the distinction between acceptance and belief, and the way that acceptance might serve as a surrogate for belief in cases where some kind of pro-attitude seems useful, see Foley (1993, pp. 22–27) (speaking of “commitment” rather than acceptance). For a similar suggestion in the context of disagreement, see Frances (2010, pp. 428–429, 450–451).

not by doxastic policies of compromise and agnosticism, but by policies of passionate advocacy and debate. Self-trust would therefore turn out to be of purely instrumental value, useful for helping us achieve the truth in the long run.<sup>27</sup>

This is a tempting line of response, but I do not think that it proves ultimately persuasive, for a number of reasons. First, to the extent there is any truth to this response, it would seem no more to justify outright belief than to justify provisional acceptance. I have just argued, however, that mere acceptance is not what we value. We value passionate, confident, original beliefs, deeply felt and tenaciously maintained. There doubtless is instrumental value in the patient, thorough exploration of the logical space of possibilities, unconstrained by whatever views currently are most fashionable. But this does not explain what value there could be in an individual's endorsing some particular position in the face of protracted controversy.

Second, the response depends on the empirical claim that self-trust does promote truth in the long run. The truth of this claim is far from obvious. After all, self-trust brings with it various epistemic liabilities. When we become convinced of a certain view, it is harder to appreciate the arguments for the other side, and there is a tendency to cut off serious inquiry. To trust oneself is to be partial, and so represents the very opposite of the impartiality that one would naturally suppose best promotes epistemic progress. Yet however these empirical facts fall out, they do not seem to account for the value we in fact place on self-trust. As things are, we have no way of knowing whether self-trust is conducive to inquiry. Even so, I have been arguing, we value it all the same. It seems unlikely, then, that we value it because we are persuaded of a certain picture of its instrumental usefulness. It is hard to see how such unknowns could be what drive our attitudes.

Finally, strictly speaking, the proposal does not save the epistemic rationality of self-trust, inasmuch as it would remain epistemically irrational for me personally to adhere to such a policy. To be sure, the present suggestion would undermine my central hypothesis, that self-trust has intrinsic, non-epistemic value. But it would do so by introducing another kind of value—the common epistemic good—which self-trust secures at the cost of one's personal interest in proportioning one's beliefs to the evidence. This is to say, then, that the present proposal itself would require sacrificing epistemic rationality, as defined in Sect. 2, and would serve as a counterexample to the thesis of evidentialism, as defined in Sect. 7. To be sure, the norm under consideration is itself epistemic, inasmuch as it remains focused on achieving truth. But by switching to the goal of long-term communal success, the response in fact lends support to my broader hypothesis that the norms governing belief extend beyond the simple epistemic requirement that I proportion my beliefs to my evidence.

<sup>27</sup> See Kelly (2005 pp. 190–192) and Elgin (2010, pp. 67–68). For a recent hypothesis regarding the adaptive value of self-trust in the context of disagreement see Mercier and Sperber (2011). The usually cited source for this line of thought is Mill, *On Liberty* ch. 3. But Mill calls only for “the fullest and freest comparison of opposite opinions” (p. 185), something that on its face seems best secured not by self-trust but by suspending belief with respect to controversial questions.

## 11 Encroachment on justification

Discussions of peer disagreement often take for granted that if a policy of giving equal weight is epistemically rational, then beliefs that contravene it will be unjustified, and so fail to count as knowledge.<sup>28</sup> Yet there is reason to suspect that in certain kinds of cases the value of self-trust is such as to justify belief even when it fails to satisfy the standards of epistemic rationality. Consider one final case:

Michael is a professional philosopher, and he believes he has solved the problem of induction. The argument, so far as he can see, is airtight, and he comes to feel practically certain of his success. But when he shows his work to his esteemed colleague, Carol, he learns that she, after careful examination, regards the argument as unsound. Michael shows the argument to a second colleague, then a third, and then a fourth, all of whom he regards as his epistemic peers. None find the argument successful. Ultimately, even though Michael himself continues to regard the argument as valid, he convinces no one else in the philosophical community.<sup>29</sup>

The equal-weight view might well seem to apply here. Indeed, given the quantity and quality of the opposition, it might seem that Michael should go beyond suspending belief, and should conclude that his argument is very likely unsound—even though he himself continues to find it intrinsically compelling. What about the influence of self-trust? Would we think it reasonable, in a broader-than-epistemic sense, for Michael to continue believing his argument? As always, we would need to hear more details about the case, including factors such as Michael's past track record of success, whether he had a clear sense of why Carol and others rejected the argument, and whether he simply retreated into his own world, or instead continued to engage his colleagues, trying to answer their objections. Let us suppose, though, that Michael's behavior in all these ways was admirable, and that he continued to believe in the argument, for the rest of his career, against the universal judgment of his peers. Then suppose, some years after his death, that the argument's soundness becomes generally recognized. Would we then say that Michael *knew* a solution to the problem of induction? It seems to me that we would be at least tempted to do so. Even if this is a case where epistemic rationality required epistemic impartiality toward his epistemic peers, the value we give to self-trust leaves us with a countervailing temptation to say that Michael was behaving admirably, and so that his belief was justified.

I do not suppose that intuitions here will be completely clear or universally shared. But my argument does not require consensus about whether it is right to say that Michael has knowledge. All I claim is that we can feel a certain pull toward that verdict, and that the proper diagnosis of the pull is that our conceptions of

<sup>28</sup> See e.g. Feldman (2006, p. 217): "My conclusion will be that, more often than we might have thought, suspension of judgement is the epistemically proper attitude. It follows that in such cases we lack reasonable belief and so, at least on standard conceptions, knowledge." See also Kornblith (2010) *passim*.

<sup>29</sup> The example is patterned on the mathematician case in Kelly (2010, p. 137), though he puts the case to different use.

justification and knowledge are not as tightly bound to epistemic rationality as we might have thought. Admittedly, those who reject the equal-weight view on the grounds that it causes us to ignore the evidence may at this point insist that Michael's case illustrates something else altogether: how it can be rational to stick to one's own evidence, even in the face of disagreement. This is, I acknowledge, one way to explain the intuition we feel about this case. But if, as has been my assumption throughout, the equal-weight view is (roughly, at least in many cases) correct as a matter of epistemic rationality, then we have another way of explaining why we are sympathetic to Michael's trusting in his evidence. For we can say that the doxastic value of self-trust infects our intuitions about such cases, making inroads into a domain where one might well expect rationality to be paramount. This runs contrary to what has been generally assumed in discussions of peer disagreement, but it should not be altogether surprising, given my argument. For inasmuch as the concept of justification itself has a normative dimension, it makes good sense that this concept might be subject to the influence of whatever doxastic norms have a bearing on the case.

One might have doubts about whether the scenario as I have described it is coherent. It requires Michael to be in a state where he understands why his epistemic peers reject the argument, and yet still he sees the soundness of the argument, but even so he remains unable to persuade his peers. It is tempting to think that this could be possible only if they are not peers at all—Michael being either a significantly better philosopher or a significantly worse one. One might make the case less incredible by imagining that Michael is able to persuade half, but only half, of his peers. Then we would have the usual sort of peer disagreement, sufficient to make suspension of belief epistemically rational. It would then be still more clear that Michael's belief might prove in time to have been justified, even in the absence of consensus among his peers. But I think the more extreme version of the case, as stated, is coherent. We need not suppose Michael takes himself to have a demonstrative proof that goes down to self-evident first principles. Justification can be had with less than that, and the point of insisting that Michael's argument was eventually vindicated by common consent is to suggest that it was an argument with sufficient merit to justify belief. Hence there can be sound, justifying arguments that do not compel consent from every informed listener. Still, one might wonder, if Michael and his colleagues were peers, how could he see the argument's merits while they did not? Here it is critical to bear in mind that Michael need have only the reasonable expectation that his colleagues are his peers. Posterity might judge Michael to have been a transcendent genius who was able to grasp what no one else could. Still, he might have no good reason to see himself in that lofty way, and it is how he is entitled to see the situation that matters, as far as epistemic rationality is concerned.

The case of Michael might suggest that skepticism is being avoided by adapting an externalist approach to knowledge—Michael can know, the thought would be, provided that in fact his methods were reliable, whether or not he was in a good position to recognize their reliability. But this is not how externalism would ordinarily be applied here. Although Michael's method of reasoning about induction was evidently reliable, his overall belief practices turned on his ignoring the

overwhelming disagreement of his peers. Reliabilists would therefore be expected to deny that his belief is justified, either on the grounds that his overall method was unreliable, or on the grounds that the peer disagreement defeated the local reliability of his reasoning about induction. The reason we are inclined to regard Michael as justified in his belief is fundamentally an internalist one—it is that we think he is doing it right or, at any rate, right enough, inasmuch as he is admirably displaying self-trust in the face of daunting odds.<sup>30</sup>

## 12 Conclusion

The puzzle of peer disagreement trades on a confusion of sympathies. On the one hand, we admire those who pursue a rigorous policy of epistemic rationality, withholding belief where the evidence is insufficient. On the other hand, we also admire those who courageously battle against long odds, defending their beliefs even when they are unlikely to be right. For as long as one associates justification with rationality, and understands rationality in its narrow epistemic sense, the equal-weight view will make trouble for knowledge. But once we untangle these various threads, we can make sense of our seemingly irreconcilable intuitions. Because we value self-trust, we are licensed at least in certain cases to set aside a policy of impartiality, and hold firm to our convictions. If those convictions turn out to be true, and well supported, then perhaps we can even rightly claim to have knowledge. Hence the equal-weight view need not lead to skepticism.

**Acknowledgment** Thanks for their advice to Michael Huemer, David Christensen, Bradley Monton, Graham Oddie, and particularly to the critical comments of a series of anonymous referees who, by undermining my own self-trust, made this a vastly better paper.

## References

- Adams, R. M. (1987). The virtue of faith. *The virtue of faith and other essays in philosophical theology* (pp. 9–24). Oxford: Oxford University Press.
- Adler, J. E. (2002). *Beliefs own ethics*. Cambridge: MIT Press.
- Baier, A. (1980). Secular faith. *Canadian Journal of Philosophy*, 10, 131–148.
- Chignell, A. (2010). The ethics of belief. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Fall 2010 Edition). <http://plato.stanford.edu/archives/fall2010/entries/ethics-belief/>.
- Christensen, D. (2007). Epistemology of disagreement: The good news. *Philosophical Review*, 116, 187–217.
- Cohen, S. (2013). A defense of the (almost) equal weight view. In D. Christensen & J. Lackey (Eds.), *The epistemology of disagreement: New essays* (pp. 98–119). Oxford: Oxford University Press.

<sup>30</sup> On Lasonen-Aarnio (2010)'s intriguingly different account, someone in Michael's position could count as having knowledge for externalist reasons, in virtue of his reliable original method of belief formation. The defeating evidence of his peers would not undermine knowledge, but it would show him to be epistemically unreasonable. Her account agrees with mine in allowing knowledge to come apart from rationality in a case like this, but she pushes much further, arguing for the "radical option" (p. 2) that knowledge is in general consistent with such defeaters. This strikes me as having incredible results, however, unless it is limited to cases where the refusal to heed the defeater can be attributed to doxastic values of the kind I describe here.

- D'Arms, J., & Jacobson, D. (2014). Wrong kinds of reason and the opacity of normative force. In R. Shafer-Landau (Ed.), *Oxford studies in metaethics* (Vol. 9, pp. 215–244).
- Elga, A. (2007). Reflection and disagreement. *Nous*, *41*, 478–502.
- Elga, A. (2010). How to disagree about how to disagree. In Feldman & Warfield (Eds.), *Disagreement* (pp. 175–186). New York: Oxford University Press.
- Elgin, C. Z. (2010). Persistent disagreement. In Feldman & Warfield (Eds.), *Disagreement* (pp. 53–68). New York: Oxford University Press.
- Enoch, D. (2010). Not just a truthometer: Taking oneself seriously (but not too seriously) in cases of peer disagreement. *Mind*, *119*, 953–997.
- Feldman, R. (2006). Epistemological puzzles about disagreement. In S. Hetherington (Ed.), *Epistemology futures* (pp. 216–236). New York: Oxford University Press.
- Feldman, R. (2009). Evidentialism, higher-order evidence, and disagreement. *Episteme*, *6*, 294–312.
- Foley, R. (1987). *The theory of epistemic rationality*. Cambridge: Harvard University Press.
- Foley, R. (1993). *Working without a net: A study of egocentric epistemology*. New York: Oxford University Press.
- Foley, R. (2001). *Intellectual trust in oneself and others*. Cambridge: Cambridge University Press.
- Frances, B. (2010). The reflective epistemic renegade. *Philosophy and Phenomenological Research*, *81*, 419–463.
- Hawthorne, J., & Bovens, L. (1999). The preface, the lottery, and the logic of belief. *Mind*, *108*, 241–264.
- Hazlett, A. (2014). Entitlement and mutually recognized reasonable disagreement. *Episteme*, *11*, 1–25.
- Hills, A. (2009). Moral testimony and moral epistemology. *Ethics*, *120*, 94–127.
- Kaplan, M. (1996). *Decision theory as philosophy*. Cambridge: Cambridge University Press.
- Keller, S. (2004). Friendship and belief. *Philosophical Papers*, *33*, 329–351.
- Kelly, T. (2003). Epistemic rationality as instrumental rationality: A critique. *Philosophy and Phenomenological Research*, *66*, 612–640.
- Kelly, T. (2005). The epistemic significance of disagreement. *Oxford Studies in Epistemology*, *1*, 167–196.
- Kelly, T. (2010). Peer disagreement and higher-order evidence. In Feldman & Warfield (Eds.), *Disagreement* (pp. 111–174). Oxford: Oxford University Press.
- Kelly, T. (2011). Following the argument where it leads. *Philosophical Studies*, *154*, 105–124.
- King, N. L. (2012). Disagreement: what's the problem? Or a good peer is hard to find. *Philosophy and Phenomenological Research*, *85*, 249–272.
- Kornblith, H. (2010). Belief in the face of controversy. In Feldman & Warfield (Eds.), *Disagreement* (pp. 29–52). Oxford: Oxford University Press.
- Lackey, J. (2010). A justificationist view of disagreement's epistemic significance. In A. Haddock, A. Millar, & D. Pritchard (Eds.), *Social epistemology* (pp. 298–325). Oxford: Oxford University Press.
- Lasonen-Aarnio, M. (2010). Unreasonable knowledge. *Philosophical Perspectives*, *24*, 1–21.
- Lehrer, K. (1997). *Self-trust: A study of reason, knowledge and autonomy*. Oxford: Clarendon Press.
- Locke, J. (1975). An essay concerning human understanding. In P. H. Nidditch (Ed.), *A treatise of human nature*. Oxford: Clarendon.
- Locke, J. (1996). Some thoughts concerning education and of the conduct of the understanding. In R. W. Grant & N. Tarcov (Eds.), *Aristotle's conception of freedom*. Indianapolis: Hackett Publishing.
- McGrath, S. (2011). Skepticism about moral expertise as a puzzle for moral realism. *Journal of Philosophy*, *108*, 111–137.
- Mercier, H., & Sperber, D. (2011). Why do humans reason? Arguments for an argumentative theory. *Behavioral and Brain Sciences*, *34*, 57–74.
- Mill, J. S. (1974). In M. Warnock (Ed.), *Utilitarianism, on liberty, essay on Bentham*. New York: New American Library.
- Pasnau, R. (forthcoming-a). Veiled disagreement. *Journal of Philosophy*.
- Pasnau, R. (forthcoming-b). Snatching hope from the jaws of epistemic defeat. *Journal of the American Philosophical Association*.
- Pettit, P. (2006). When to defer to majority testimony—and when not. *Analysis*, *66*, 179–187.
- Russell, B. (1945). *A history of Western philosophy*. New York: Simon and Schuster.
- Schroeder, M. (2012a). Stakes, withholding, and pragmatic encroachment on knowledge. *Philosophical Studies*, *160*, 265–285.
- Schroeder, M. (2012b). The ubiquity of state-given reasons. *Ethics*, *122*, 457–488.
- Shah, N. (2006). A new argument for evidentialism. *Philosophical Quarterly*, *56*, 481–498.
- Sliwa, P. (2012). In defense of moral testimony. *Philosophical Studies*, *158*, 175–195.



- Sosa, E. (2007). *A virtue epistemology: Apt belief and reflective knowledge* (Vol. 1). Oxford: Clarendon Press.
- Sosa, E. (2010). The epistemology of disagreement. In A. Haddock, A. Millar, & D. Pritchard (Eds.), *Social epistemology* (pp. 278–297). Oxford: Oxford University Press.
- Steup, M. (2001). *Knowledge, truth, and duty: Essays on epistemic justification, responsibility, and virtue*. Oxford: Oxford University Press.
- Stroud, S. (2006). Epistemic partiality in friendship. *Ethics*, 116, 498–524.
- van Inwagen, P. (1996). It is wrong, everywhere, always, and for anyone, to believe anything upon insufficient evidence. In J. Jordan & D. Howard-Snyder (Eds.), *Faith, freedom, and rationality: Philosophy of religion today* (pp. 137–153). Lanham: Rowman & Littlefield.
- White, R. (2009). On treating oneself and others as thermometers. *Episteme*, 6, 233–250.