Commitment Beyond Justification

Paul Katsafanas
Boston University

Forthcoming in *Ergo*

Abstract:

Should our degree of commitment to a value, relationship, or goal be proportional to the degree of justification that we take the commitment to possess? Or are there reasons for maintaining wholehearted commitments even in cases where we have relatively weak justifications for those commitments? I argue in favor of the latter position: degree of commitment should sometimes diverge from degree of justification. To make this case, I introduce and critique what I call *Locke's Dictum:* the claim that our degree of our commitment to an evaluative claim should match the degree of justification for it (decisive justification warrants full commitment, whereas weaker justification warrants weaker commitment). While it has initial appeal, I argue that Locke's Dictum is mistaken. I argue that many important goals, values, and relationships require wholehearted commitment to evaluative propositions in the absence of decisive justification for them. So we are faced with a choice: either we violate Locke's Dictum; or we forgo the goods internal to full-fledged commitments; or we deny a reasonable pluralism about possible commitments. I argue that we must reject Locke's Dictum and with it a series of seeming platitudes about the connection between justificatory reflection, commitment, and fanaticism or extremism.

1. Locke's Dictum

John Locke tells us that one "unerring mark" of the love of truth is "the not entertaining any proposition with greater assurance, than the proofs it is built upon will warrant" (Locke 2000: Book Four, Chapter XIX, Section 1, 583). He seems to be endorsing a straightforward relationship between credence and justification, which we might describe as follows:

> **Locke's Dictum, provisional formulation**: the degree of credence that we give to a proposition should be proportional to our degree of justification for accepting it.

Locke's Dictum is most tempting and easiest to understand in the case of beliefs about straightforward factual matters. If I have equal justification for thinking that it's raining and that it's not raining, I should withhold assent from both propositions rather than giving full credence to either. Or suppose we take something more complicated. I try to figure out how much global warming will occur by 2050. Some models put it at 2 degrees, others at 4, others at 5, whereas others give still different answers. If I give full credence to the proposition that there will be exactly 4 degrees of warming by 2050, I've again outstripped my justification. I've given more credit to this

proposition than the principles and proofs would support; what they actually support is some qualified credence in the claim.

So, although Locke's Dictum is perhaps controversial, we can at least see what it would mean in the case of beliefs about factual matters. If you don't have decisive justification for some propositions about factual matters, you should adjust your credences. Credence should track justification. This seems like an intuitive position to hold.[1]

But the claim I've quoted occurs in an essay on *enthusiasm*, which is what we would today call fanaticism, extremism, or zealotry. The enthusiast, as Locke is conceiving him, is an "untractable zealot" who takes himself to be an unerring authority on matters of practical import (Locke 2000: Book IV, Chapter XIX, Section 11, 588). The enthusiast may cite divine revelation or immediate insight into reality or religious dogma, but what's important for our purposes is that the "propositions" that he gives absolute certainty are evaluative or practical. To use one of Locke's examples, the enthusiastic Quaker thinks that he can compel others to adopt his religious practices.[2] Or, to use an example that Locke was familiar with: Voltaire claims that

> The most detestable example of fanaticism was that of the burghers of Paris who on St. Bartholomew's Night went about assassinating and butchering all their fellow citizens who did not go to mass, throwing them out of windows, cutting them in pieces. (Voltaire, *Philosophical Dictionary,* "Fanaticism")

Here, Voltaire is referring to a massacre of French Protestants by French Catholics, carried out on Aug 24-25, 1572. The initial massacre left approximately 3000 dead and many more injured; repeated attacks took place throughout France over the next several months and led to thousands of additional deaths and injuries. Or consider another example from the period: in 1534-5, a group of radical Anabaptists overthrew the government of Münster and briefly installed a theocratic commune led by John of Leiden, who claimed that visions from God justified his absolute rule of "new Jerusalem." This led to a protracted siege of the city, mass starvation, and many deaths in battle. These are the sorts of events that Locke is contemplating.

So the fanatic is not just entertaining propositions about rain and temperature. He is entertaining and granting authority to propositions such as—and again, these examples from Locke's context—"non-Quakers should be persecuted," "I ought to slaughter Protestants," "we should violently overthrow the government and install a theocracy," and so on. Those kinds of commitments carry grave consequences: they rationalize violence, coercion, and oppression. We can see why Locke wants them to be controlled by critical reflection.

But how should we understand Locke's Dictum in that context? Notice that it's not just about the *acceptance* of propositions. The fanatic and the ordinary Catholic believer might *share* the belief that

---

[1] At least, it seems like an intuitive position to hold about *believing* evaluative propositions. But belief is not the only way of accepting propositions (cf. Jackson 2021). Some accounts of faith (e.g. Buchak 2017a) and hope (e.g. Calhoun 2018 and Martin 2013) treat these notions as having a weaker epistemic component than belief. So even if belief in a proposition requires a high degree of justification or evidence, faith or hope in that proposition might be compatible with a lesser degree of justification or evidence. While I lack the space to address hope, I discuss faith in Section 6.

[2] "I am weary of these Quakers" (Locke 1978: 44). Locke mentioned the Quakers by name in a draft, but removed the reference in the published version. For a thorough defense of the idea that Locke had the Quakers specifically in mind when writing this essay, see Casson 2024: 86-88 and Anstey 2019.

Protestants embrace the wrong faith; both might be absolutely certain about this belief; but the fanatic takes this as a reason for persecuting Protestants and the ordinary believer doesn't. Or, the fanatic and the ordinary believer might share the belief that theocracy ought to replace the feudal system; but the fanatic joins a peasant rebellion and the ordinary believer doesn't.

The difference between the fanatic and the ordinary believer needn't show up merely in the contents of these evaluative beliefs.[3] It shows up in the extent to which they let these beliefs override competing values and, more generally, in the degree of self-sacrifice and other-sacrifice that they are willing to engage in so as to support the evaluative belief. This is why, immediately after the quotation with which I began this essay, Locke goes on to say:

> Whatsoever credit or authority we give to any proposition more than it receives from the principles and proofs it supports itself upon, is owing to our inclinations that way, and is so far a derogation from the love of truth as such. (Locke 2000: Book IV, Chapter XIX, Section 1, 583)

Here, Locke speaks not just of the "entertaining" or accepting of propositions, but of the *authority* that we give to them: we treat these propositions as having *greater weight* than competing evaluative propositions and we accept sacrifices so as to act in accordance with these propositions.

In short, and in a sense that I will explore more fully below: Locke is interested not just in how theoretical acceptance of an evaluative proposition relates to justification, but in how *granting authority* to an evaluative proposition relates to action, motivation, and justification. I will describe this in terms of *commitment,* where commitment is understood to involve both the acceptance of a proposition and the acquisition of some tendency to live in accordance with it. Below, I will explain commitment in more depth, but for now we can put Locke's point as follows: even when they believe the same propositions, the fanatic and the ordinary believer differ in the strength of their commitment to these propositions. And Locke thinks that we can formulate a claim about how commitments and justifications relate:

> **Locke's Dictum**: our degree of commitment to an evaluative proposition should be proportional to our degree of justification for accepting the proposition.

~ ~ ~

Now, in a way the beginning of this paper has been misleading. For I'm not really interested in Locke interpretation. What I am interested in, and what I want to discuss, is the claim that I'm calling Locke's Dictum. I think Locke is picking up on a deep tension in our thinking about the relationship between commitment and justification. My main goal in this paper is to bring this tension into view. I'd *like* to resolve the tension, and at the end I will make some movements in that direction. But I don't think the tension is fully resolvable and in any case my main aim is to get it into view.

[3] Of course, the difference might show up in *other* evaluative beliefs. For example, one person might believe that persecution is warranted and another might not. Without denying that some cases will be explained in that way, I want to explore a different way of accounting for the divergence between the fanatic and the ordinary believer. For accounts of fanaticism, see for example Cassam 2022, Battaly 2023, Katsafanas 2023a, and Katsafanas 2023b.

As a first step, let me draw attention to the way that commitments can be flawed or problematic. One problem is that commitments are too weak. Something that merits strong commitment gets only weak commitment, as when the lackadaisical parent neglects his child or when we fail to give appropriate concern to climate change. Another problem is that commitments are too strong. Something that merits only a tentative or weak commitment gets a wholehearted one, as when the fanatic remains steadfastly committed to highly questionable goals, or when a person with the most dubious of religious views steadfastly and resolutely commits himself to imposing these views on others.[4]

So commitments can be too weak or too strong. And Locke's Dictum tries to address this by linking *degree of commitment* to *degree of justification*. It says: apportion your degree of commitment to your degree of justification for believing that commitment.

That idea has some intuitive force. After all, if you're not certain that a particular commitment is justifiable, it looks like you should be more hesitant with respect to acting on it, especially in cases that would involve overriding other values or imposing costs on yourself or others. If you're going to join a peasant rebellion or persecute members of a different faith, you'd better be *very* certain that the relevant evaluative propositions are justifiable.

So Locke's Dictum might look appealing. It certainly looks appealing as an antidote to fanaticism and extremism. But, more generally, it might seem to posit an exemplary relationship between commitment and justification, one that all of us would do well to heed. However, I will argue that Locke's Dictum is flawed and should not be accepted. It provides a distorted and damaging picture of the relationship between commitment and justification.

I begin, in Section 2, by clarifying Locke's Dictum. With that groundwork in place, Section 3 examines how our notions of integrity bear on Locke's Dictum. Section 4 argues that given a reasonable degree of epistemic humility, Locke's Dictum precludes integrity; it thus has significant costs. Section 5 argues that these costs are even higher. It examines the role that strong (or "resolute") commitments play in social and ethical life, arguing that Locke's Dictum would undermine them and that giving them up is unimaginable. Locke's Dictum, though it rules out fanaticism and extremism, is thus problematic. Section 6 asks whether we can resolve this problem by disentangling degree of commitment from degree of justification. There, I offer some preliminary reflections on how Locke's Dictum might be reconfigured so as to rule out the problematic cases without also precluding resolute commitment more generally. Section 7 concludes.

2. Clarifying Locke's Dictum

Locke's Dictum attempts to link two items: justifications and commitments to evaluative claims. Let's make this more precise by introducing some distinctions and clarifications.

---

[4] Commitment's strength can pick out either its lexical ranking (how many thing it does or should override) or its motivational force (holding the lexical status constant, whether the agent's motivational engagement with the object track this lexical status). More on this in Section 2.3, below.

2.1 The Nature of Commitment

We will start with the nature of commitment. What is a commitment? To be committed to some object is to exhibit a temporally extended relationship to it, where this relationship involves a cluster of intentional and affective states.[5] In general, we speak of commitment when we are considering relatively long-term goals, activities, or relationships that require some degree of perseverance in order to be achieved or maintained successfully. It would be odd to say that I am committed to getting a cup of coffee if all that this requires is effortlessly walking to my kitchen; it would be somewhat more natural to speak of a commitment to getting coffee if this requires a more arduous, long-term process, such as walking through a snowstorm to the local coffee shop. So commitment seems to involve a willingness to pursue an end despite obstacles. These obstacles can take many different forms: difficulties, boredom, disengagement, distraction, objections, the recognition of other appealing options, and so on.

Commitment comes in degrees. If we compare my commitment to being on time for my office hours to my commitment to promoting my child's welfare, the latter is *much* more robust. Although I am committed to being on time, I can imagine many circumstances in which I would abandon this commitment: if the weather is really disastrous, or if I have a cold, or even if I run into a friend who I haven't seen in many years, I might be a bit late for my office hours or even cancel them. Not so with the commitment to my child's welfare: I wouldn't imperil her welfare on these grounds.

Being committed to some end differs from intending that end. They share some features: intentions also involve some degree of willingness to withstand obstacles and contrary pressures (if I claim that I intend to run a marathon, but give up at the first hint of difficulty or the first temptation to do something else, we might be skeptical of the claim that I genuinely intended to run the marathon—perhaps I merely wished or desired to do so).[6] But commitments are more robust than typical intentions.

How so? On the account that I favor, commitments involve both an intention *and* a disposition to maintain the intention.[7] Return to my example of fetching a cup of coffee. Suppose I start to head out for a cup of coffee, put on my boots and coat, make it out the door, but then notice that it's really very chilly indeed; I reconsider, and decide to stay at home. I did intend to get a cup of coffee, but I reconsidered that intention and abandoned it. In a case like this, we can say that I have an intention but lack a disposition to ensure that the intention persists. I hence lack a commitment to getting the coffee.

Of course, the disposition to sustain the intention comes in varying strengths: I am willing to undertake some difficulties in order to sustain my intention to hold my office hours on time, but I am willing to undertake many more in order to sustain my intention to promote my child's welfare.

---

[5] Commitment can take many different objects, which include ends, goals, values, ideals, persons, relationships, projects, and so on. I will focus on commitment to evaluative propositions (more on this in a moment), though I think my points carry over (with some rephrasings) to the other objects of commitment.
[6] See for example Bratman 2012.
[7] This account is developed by Calhoun 2018.

In short: we can understand commitments as intentions coupled with a disposition to sustain the intention, where this disposition comes in degrees.[8]

2.2 The Relevant Type of Case

The question that Locke's Dictum raises is this: commitments come in different strengths and justifications come in different degrees. How should these relate? Locke says: they should be proportional.

This is an intuitively appealing idea. It is implicated in some of the ways we think about idealized agents (more on this soon). But it's nonetheless mistaken.

To begin seeing why it is mistaken, let's spell out the relevant type of case in more detail. I want to focus on cases that have two important features: (1) people can reasonably disagree about whether to accept an evaluative claim E1 or some competing evaluative claim E2; (2) people do not apportion their degree of commitment to E1 to the degree of justification for E1, where justification is here understood as *subjectively available* justification.

Start with (1). We want to consider hard cases, cases in which reasonable people could disagree about whether to accept evaluative claim E1 or evaluative claim E2 (or, somewhat differently, whether to accept that E1 outweighs E2). In Locke's day, the typical examples were religious beliefs, disputes about whether there should be feudal aristocracies or theocracies, disputes about whether peasant rebellions are legitimate, and so forth. But those issues and those assumptions are distant from most of us today—thankfully, only a small minority think that violent revolts in favor of theocratic or religious rule are within the bounds of reasonable disagreement. So let's consider a more contemporary example. I think reasonable people can disagree about whether and to what extent animal suffering matters. Some people think that all animal suffering *matters equally*. Some people think that all animal suffering *matters, but in different degrees*. Some people think that suffering matters only for animals at a certain level of cognitive sophistication. And others, still, think that suffering only matters for self-conscious animals. And to these views we can add nuances: perhaps the extent to which suffering matters comes in degrees, so that the badness of a pig's suffering is lexically outweighed by any degree of human suffering. So we could have: E1 = all animal suffering matters equally; E2 = suffering matters only in the case of cognitively sophisticated animals. (If you don't like this example, feel free to substitute another; nothing turns on the particulars.)

So the first point is that we are considering cases in which, given certain starting assumptions, reasonable agents can disagree about which evaluative claims to accept. Moving on to claim (2): Locke is claiming that some individuals—the "enthusiasts" or fanatics—do not apportion their degree of commitment to the results of justificatory reflection. Despite lacking adequate grounds for thinking that E1 is correct and E2 incorrect, they wholeheartedly commit themselves to E1.[9]

---

[8] While I adopt Calhoun's account of commitment, nothing in this essay turns on the details of her account. Any account that introduces a distinction between ordinary intention and more robust forms of commitment would be compatible with my argument. For example: Holton 2009 and Andreou 2022 provide insightful analyses of "resoluteness," which in some ways mirrors Calhoun's notion of commitment. Bratman 2012 and 2018 argues for a concept of planning agency according to which it can be rational for an agent to adhere to a plan even if she finds that her present evaluative perspective warrants abandoning it. See also Chang 2013 on willing as a form of commitment.
[9] Wholeheartedness has been analyzed in different ways; for two influential examples, see Frankfurt 2004 and Calhoun 1995. I am using the term "wholehearted commitment" to pick out commitments that are both strong and non-

When Locke speaks of justification, he's interested in what a person is in a position to know, rather than what a person actually knows. Not everyone reflects on their justifications for their evaluative commitments; and when people do reflect, they sometimes make mistakes. So Locke has the following case in mind: based on the evidence that she has available, the fanatic *is in a position to know that her commitment is not adequately justifiable,* yet her commitment is strong.[10]

That's the sort of case I want to consider. The fanatic may tell himself that he has some mystical insight into E1's justification, some revelation from God, or whatever: but the fanatic is in a position to know that he lacks certainty. There may be *some* grounds for adopting E1, but there are not decisive grounds. So we are considering cases in which reasonable people are in a position to know that they lack full justification for the evaluative proposition to which they are committed.

To review, then, we will focus on cases in which: (1) people can reasonably disagree about whether to accept evaluative claim E1 or some competing evaluative claim E2; (2) people do not apportion their degree of commitment to E1 to the degree of justification for E1, where justification is here understood as *subjectively available* justification.

2.3 Justificatory Status of Commitments

Locke's Dictum says that your degree of commitment to an evaluative proposition should be proportional to its justificatory stranding. We can illustrate this with a version of Locke's example. Suppose you think that one Christian sect is blaspheming and is hence condemning itself to eternal damnation. If this were true, and if you had decisive justification for the belief that it were true, then it could make sense to forcibly convert these blasphemers. After all, whatever worldly sufferings you inflict on them will pale in comparison to eternal damnation and will be redeemed by eternal salvation. But what if you don't have *decisive* proof? What if you think there are good, but not fully convincing, reasons for thinking that the sect is condemning itself to damnation? At the very least, you might think that you should attenuate your commitments. Violent conversion would demand a very high degree of certainty; anything less should lead you to rethink your project. Perhaps, instead of violent coercion, all that's warranted is peaceful attempts at conversion.

If we consider these sorts of cases, it seems like the practical analogue of partial credence would be deliberative and motivational attenuation. The agent who follows Locke's Dictum would be less strongly disposed to pursue (some of) the practical implications of her evaluative beliefs, particularly when doing so clashes with the practical implications of other possible evaluative beliefs.

---

ambivalent. A wholehearted commitment will be strong in the sense that it overrides competing commitments in practical deliberation and in action; and it will be non-ambivalent in the sense that the agent unwaveringly acts in accordance with it. I explain this in more detail in Section 2.3.

[10] I interpret Locke as noted above because in Book Four, Chapter XIX, Sections 10-11, Locke says that enthusiasts take themselves to have divine revelation or inspiration from God, but, were they to ask themselves whether they can distinguish these putative revelations from "perception of an inclination or fancy", they would see that they lack grounds for doing so. "But however it be called light or seeing, I suppose it is at most but belief and assurance: and the propositions taken for a revelation is not such as they know to be true, but take to be true. For where a proposition is known to be true, revelation is needless: and it is hard to conceive how there can be a revelation to any one of what he knows already. If therefore it be a propositions which they are persuaded, but do not know, to be true, whatever they may call it, it is not seeing, but believing" (Section 10). Or, more simply: "Enthusiasm fails of evidence, that the proposition is from God" (Section 11).

Because commitments have a role in both deliberation and action, the degree of a commitment could refer to at least two distinct things:

(A) Deliberation: the weight or lexical status of the commitment's object (how many other things it overrides within practical deliberation).
(B) Action: the degree of motivational investment the agent has in the commitment (how motivated she is to abide by it, to set aside lexically lower ranked goods in favor of it, etc.).

On (A), decreasing the strength of a commitment would mean adjusting the extent to which you take it to override other values within practical deliberation. To use one of Locke's examples, you might go from thinking that religious values should override all other values (tolerance, harmonious coexistence, etc.) to thinking that religious commitments should not override these other values. Or, in the animal suffering case, you might go from thinking that economic considerations must be sacrificed to animal welfare to thinking that there is no particular weighting of these considerations that must be maintained. What formerly functioned as overriding within practical deliberation would start to be treated as fungible.

On (B), decreasing the strength of a commitment would mean preserving its weight or lexical status while varying the actions that you are willing to take so as to secure it. So, you might hold on to the belief that Catholics should be forcibly converted but decline to act on it. Or, you hold on to your belief that animal suffering matters, but you no longer strive to prevent economic development when it increases animal suffering. In a way, this would look like a form of akrasia—but in this case, it would be a deliberate form of akrasia, perhaps better called forbearance. So, this is a case in which you'd think that animal suffering matters; but you wouldn't be totally confident that you can justify that proposition; so you'd hesitate to act on it when it imposes costs.

Notice that on *either* of these strategies, your commitment would be weakened. Commitments have two aspects: they involve a particular weighting of normative entities and a willingness to perform a particular set of actions. When we speak of a person's being strongly or wholeheartedly committed to an evaluative claim, we can mean either or both of these things. For example, if I say that Ben is wholeheartedly committed to the alleviation of animal suffering, I can mean either that a lexical ranking of his values would place this quite high or that he is strongly and resiliently disposed to act in ways that promote that value. In the cases that we're considering, either or both of these will occur; and in that sense, the commitment would be weakened.

So *weakening a* commitment could mean either decreasing its weight within deliberation or decreasing the extent to which the agent strives to realize it (or both). Locke's Dictum tells us to weaken commitments in both of these ways when they lack decisive justification.

3. Fanatics, Extremists, and People of Integrity

Assume we accept Locke's Dictum. It seems to give us a picture of the ideal agent. The ideal agent would be the *person of integrity*. That is, the ideal agent would be the person who is subjectively certain about the justificatory standing of her commitments and has extremely strong commitments.

The notion of integrity plays different roles within philosophy, but one dominant notion is that the person of integrity is distinguished by the fact that she resolutely maintains her commitments despite challenges, costs, and temptations.[11] Standard examples from the philosophical literature include Socrates; Antigone; Thomas More; Martin Luther; Gandhi.

We can detect two dimensions in these portraits of integrity. First, there is a *practical* dimension: these people are wholeheartedly and unwaveringly committed to some end (or activity, ideal, value, relationship, etc.). They persevere despite challenges. They are willing to make immense sacrifices (of both themselves and others) for the end.

So, when Martin Luther comes to believe that the Catholic Church is corrupt and mistaken, he risks excommunication and death rather than concealing or compromising these beliefs. When Antigone sees that Creon will have her killed for burying her brother Polynices, she nonetheless refuses to compromise, citing the superiority of divinely ordained burial rites over human law. Or, when Bernard Williams' version of Gauguin sees that he can maintain and realize his artistic commitments only by abandoning his family, he does so (Williams 1981). And examples could be multiplied. But the core idea is straightforward: these agents persevere, maintaining their commitments even when doing so involves great costs.

That's the practical side. But, second, there is an *epistemic* dimension. When these people engage in justificatory reflection about their goals and values, they see them as warranted. Martin Luther engages in rigorous, lifelong theological reflection; Antigone has no doubt that divine law requires her compliance. These individuals do not blindly or unreflectively commit themselves to what they value: rather, they commit themselves wholeheartedly because they take themselves to have reasonable certainty that the thing to which they're committing themselves merits this commitment. And, though I'll raise some doubts about this in the next section, let's assume for now that this subjective certainty is adequately justified: given what these people are in a position to know, they justifiably regard their ends as worthy of total commitment.

In the ideal case of integrity, the person would be wholeheartedly and unwaveringly committed to an object only if she *justifiably* regards this object as worthy of total commitment. And that's just what these individuals seem to exhibit: they are subjectively certain that their (extreme) degree of commitment to their values is justifiable.

So this looks like an exemplary case of fulfilling Locke's Dictum. And not only that. Additionally, these cases of integrity can be used to *support* Locke's Dictum. For consider what happens when we imagine varying certain aspects of a person of integrity.

If you preserve the degree of subjective certainty but weaken the degree of commitment, the person looks weak-willed or at least less exemplary. Our esteem fades. There's nothing especially admirable about a person who takes wholehearted commitment to some cause to be justifiable, but who can't muster the motivational energies to manifest that commitment.

---

[11] "The person of integrity stands by his commitments and convictions" (Svavarsdottir 2015: 62). "The person possessing integrity is true to his commitments" (Taylor 1981). "We use the term [person of integrity] for someone who lives up to his own standards. And that is because we think that living up to them is what makes him one" (Korsgaard 1996: 102).

If you weaken the degree of subjective justification but preserve the degree of commitment, the person looks like a fanatic or dogmatist. Again, our esteem fades. We can think of plenty of cases in which a person is in a position to entertain doubts about some end but nevertheless commits to it wholeheartedly. Here it's important to remember that subjective certainty refers not to the actual degree of certainty that the person has, but to the degree of certainty that she's in a position to have. The fanatic might *in fact* believe that she is justifiably certain about her end. But, so long as she's in a position to know that her certainty is untenable, she counts as lacking the relevant degree of subjective certainty. For example, a member of the Westboro Baptist Church or ISIS or Stormfront might in fact take herself to have watertight justification for some pernicious evaluative belief; but she is, let's say, in a position to know that she lacks this justification.[12] A person who lacks subjective certainty about her intolerant, pernicious commitment but nonetheless pursues it wholeheartedly looks like a fanatic. So again, this provides some support for Locke's Dictum: a person of integrity minus the subjective justification just looks like a fanatic.[13]

4. Skepticism about the Possibility of Adequate Justification

People of integrity combine wholehearted commitment with subjective certainty. But certainty is a very high bar: usually, when we are appropriately attentive to the evidence available to us and reason correctly from that evidence, we fall short of certainty. A reasonably thoughtful individual can usually find some grounds for doubting her evaluative commitments (both in particular cases and in general). She can envision reasonable alternatives to these commitments (or reasonable reinterpretations of the same commitments). And even when we're confident in some evaluative proposition, difficulties can arise. For example, we can be confident that a particular value while lacking confidence about the weight or lexical rank of this value (e.g., I can be confident that animal suffering has disvalue, but uncertain about how to weight this against considerations of justice, autonomy, and so on). We are often incapable of fully answering reasonable skeptical challenges to our weightings or rankings of values, principles, and commitments. We can see this in several ways. I will mention two.

First, consider the complexity of applying norms or values to concrete situations. Any situation that we face can be redescribed or reconceptualized; any situation is open to further articulations and revisions. So we can wonder which set of evaluative claims is applicable and which set of claims is to be prioritized *in a given case*.

Second, consider a reasonable form of pluralism. There is a plurality of reasonable yet incompatible and irreconcilable comprehensive ethical, evaluative, and philosophical views, and these yield different claims both about what has value and what the weights or lexical ranks of various values are. So we can wonder which sets of values should be prioritized in general.[14]

---

[12] For example, these people are in a position to access a certain kind of higher-order evidence, namely that their beliefs conflict with the beliefs held by others. Of course, they may dismiss these outsiders as being epistemically impaired, morally deluded, and so on. But, insofar as the fanatic is impervious to this higher-order evidence, she may count as lacking subjective justification. See DiPaolo 2020 for related discussions.

[13] For discussions of the way that integrity relates to fanaticism, see Halfon 1989, McFall 1992, Hebert 2002, Scherkoske 2013, and Svavarsdóttir 2022.

[14] See for example Wong 2006.

Although I won't defend these claims, I think most people with some degree of epistemic modesty accept versions of them. Suppose they are right. Then we get the following conclusion: in an important range of cases, there is reasonable disagreement both about whether we should be committed to E1 at all and, even more frequently, whether E1 should be weighted more heavily than E2.[15]

Suppose we accept Locke's Dictum, thinking that we should apportion our degree of commitment to justificatory standing. Then it looks like insofar as we see full-fledged commitment to E1 over E2 as questionable (or even just less than certain), we should attenuate that commitment.

So Locke's Dictum coupled with a healthy degree of epistemic modesty entails that we should rarely (perhaps never) have wholehearted commitments. For those commitments would demand a degree of certainty that we rarely (perhaps never) possess.[16]

This shows up in ordinary life. The people who are most confident about their evaluative commitments often look like *fanatics* or *dogmatists* or *true believers:* they fail to see that their evaluative beliefs are uncertain. Even if they engage in great quantities of reflective activity, justificatory reasoning, and so on, they fail to question certain premises and assumptions. (Yeats, *The Second Coming:* "The best lack all conviction, while the worst / Are full of passionate intensity")

And, in fact, this is true even of philosophers' favorite exemplars of integrity. In the previous section, I asked the reader to grant that Luther, Antigone, and the other exemplars were responding appropriately to the evidence available to them and hence counted as subjectively justified. But now I want to question that assumption. The Martin Luther who proudly proclaims "here I stand, I can do no other!" and is held up as a portrait of integrity is also the person who urges us to "smite, slay and stab, secretly or openly" the "poisonous, hurtful," and "devilish" peasants, the "mad dogs" who contravene Luther's own evaluative beliefs (Luther 1525: 55). There is no doubt that Luther takes his beliefs to be justified. But *should he?* Imagine that, were Luther appropriately sensitive to evidence and reasons, he would modify or at least question some of his beliefs. That might seem like a good thing: when we consider the costs that these confident evaluative beliefs impose, we might hope for more critical reflection, less certainty, less confidence. As these examples indicate, the people who are most confident about their evaluative commitments often look like *fanatics* or *dogmatists:* they fail to see that their evaluative beliefs are uncertain. Despite engaging in great quantities of reflective activity, justificatory reasoning, and so on, they fail to question certain premises and assumptions.

These reflections can lend some additional support to Locke's Dictum. Maybe it really is only the dogmatic, fanatical individuals who can manifest integrity and wholeheartedness. After all, resolute, wholehearted commitments look problematic—they spark discord, strife, conflict. Locke's Dictum would have us weaken these commitments (this is his avowed intent in his essay). And perhaps that's good.

---

[15] I defend this claim in Katsafanas 2023a.

[16] Strictly speaking, the argument against Locke's Dictum only requires that there be *some* cases in which we think wholehearted commitments are appropriate even in the absence of certainty. So, even those who think that epistemically modest agents regularly attain certainty can agree that Locke's Dictum will inappropriately rule out some cases of wholehearted commitment.

So here's where we are left: the kind of wholehearted, integrated commitments that Luther manifests might seem possible only if we bear an unjustifiable certainty about our commitments. So, if we accept Locke's Dictum, we would have to choose between wholehearted yet dogmatic commitments or etiolated yet reasonable ones. In short: in ruling out fanaticism, Locke's Dictum seems to rule out all forms of wholehearted commitment. And that might not seem so bad when we consider the costs of wholehearted commitment.

But let's look at this from another side. For we shouldn't forget that wholehearted commitments are implicated in some of our most valued forms of life.

5. Resolute Commitments

If Locke's Dictum plus reasonable skepticism ruled out integrity, that would be regrettable but perhaps offset by the decrease in social discord that might result. But what if the problem extends even further?

I think it does extend further. Some central, valuable features of human life demand what I will call "resolute" commitments. Resolute commitments are commitments that are strong in both of the senses from §2.3: deliberation and action. That is, resolute commitments have significant weight or high lexical status (they override or overrule other considerations within practical deliberation) and the agent is disposed to conform to these weights or lexical statuses in action, even when doing so involves sacrificing competing goods or enduring other hardships. These are just the kinds of commitments that, I've suggested, Locke's Dictum rules out.

Imagine a world in which no one had resolute commitments, in which all values were fungible and all motivation to abide by them weakened. What would that look like?

There would be practical problems. Without stable weightings of goods and perseverance in realizing the weightier goods, certain goals would be unlikely to be achieved. For example: perhaps you need resolute commitments to fighting climate change in order for this to happen. After all, we lack decisive evidence that any particular catastrophic scenario will occur. We can speculate that technological innovations will render current efforts unnecessary. We can raise doubts about the accuracy of climate models. For all of these reasons, we fall considerably short of full certainty that catastrophic changes will occur. Suppose that in light of this, we adjust our commitment to fighting climate change. We take only halfhearted, easily overridden attempts to forestall these changes. The results could be disastrous. In cases like this, there are pragmatic reasons to adopt resolute commitments.[17]

[17] And this problem might extend not just to long-term difficult activities, but to all intentional actions. Consider Bratman's influential work on intention, which argues that part of the functional role of intentions is resisting reconsideration: he contends that cognitively limited creatures have reason to adopt settled policies or plans rather than constantly reopening deliberation whenever they encounter new evidence. After all, for a cognitively limited creature reopening deliberation has costs. If this is right, then once we commit to something (or intend something) we will have some reason to maintain that commitment even when we encounter contrary evidence, and in that sense we have some reason to violate Locke's Dictum (Bratman 2012 and 2018; cf. Holton 2009, Rioux 2022).

But those cases don't really threaten Locke's Dictum in a substantial way—they just show that there can be pragmatic or instrumental reasons for abandoning Locke's Dictum in certain cases. A deeper objection to Locke's Dictum would have to be internal: it would have to show that even independently of these instrumental reasons, abiding by Locke's Dictum is problematic.

I think there is such an objection: Locke's Dictum precludes certain forms of excellence. This isn't because resolute commitment is necessary to bring about some further end, which would be conceptualizable as distinct from the committed process. It's because the resolutely committed activity is itself the good.

To illustrate this, imagine a person who is committed to his partner just to the extent that he takes this commitment to be justifiable. Countervailing considerations can arise in any relationship. For example, philosophers who write about love have sometimes thought that a puzzle exists: love seems to be based on the beloved having certain qualities (she's funny, smart, engaging); but other people might have the very same qualities to a greater extent (this person is funnier, smarter, more engaging); so, these philosophers ask, what justifies sticking with the original person rather than "trading up" to the new person? This is sometimes called the trading-up objection.[18] Now it's clear that we *can* treat our relationships this way: anyone who has engaged in casual dating knows that it's possible to commit to someone only insofar as and to the extent that no one better comes along. And according to Locke's Dictum, plus a certain view of what justifies loving relationships, this is exactly the right thing to do: your commitment should be sensitive to the degree to which it's justified. If it's justified on the basis of the qualities of the beloved, and someone with a greater degree of those qualities comes along, it seems to make sense to trade up.[19] But notice that there's a real and significant difference between envisioning one's relationship in *that* way and forming a resolute commitment to one's beloved. The resolute commitment outstrips the justification for it; it violates Locke's Dictum. And yet the resolute commitment seems praiseworthy, perhaps even a constituent of the good life.

The relationship case shows that resolutely committed activities and non-resolutely committed activities can differ in profound ways even when they are directed at the same object. The resolutely committed relationship is not the same thing as the non-resolutely committed one. Nor does it instantiate the same goods. Accordingly, a person who, abiding by Locke's Dictum, abandons resoluteness thereby deprives himself of a significant good.

And this extends beyond personal relationships. Consider a political or social cause. I mentioned, earlier, that various normative claims pertaining to animal suffering are contentious: some people think that all animal suffering matters, some think that none does, some think that different animals matter to different degrees, and so on. Suppose Frances believes that the suffering of all cognitively sophisticated animals matters. She thinks this view is defensible yet not immune to challenges; in that sense, her justification for the view falls far short of certainty. Yet she commits herself to the view resolutely, prioritizing it over competing ends and alternative evaluative views. She thereby

[18] There is a large literature on this point. See for example Howard 2019, Kolodny 2003, Protasi 2016, and Velleman 1999.

[19] Of course, we can object on several grounds: is love best conceived as based on the qualities of the beloved? Are those qualities comparable across persons? Are they degreed? Are the reasons for initiating a relationship the same as the reasons for continuing or ending a relationship? And so on. While I think the trading-up objection is misguided, I use it here only for illustrative purposes.

violates Locke's Dictum. Again, this seems praiseworthy: even independently of what it achieves, the resolutely committed life seems admirable in ways that a slack, etiolated commitment to the same evaluative view would not.

For another example—and perhaps a more contentious one—consider Bernard Williams' Gauguin case (Williams 1981). Gauguin is committed to painting in innovative ways. As Williams describes the case, Gauguin decides that he has only two options. He can stay in Paris with his wife and children, but if he does so he will have to sacrifice his goal of artistic greatness, for the Parisian life makes it impossible to attain anything more than mediocrity; or he can remain resolutely committed to his goal of artistic greatness, but this requires abandoning his wife and children and setting off for a new life in Tahiti. He chooses artistic greatness: he sacrifices his relationship and his obligations, and imposes immense costs on his wife and children, so as to remain resolutely committed to his artistic project. The general form of this case is familiar: many brilliant artists, musicians, writers, and other creative figures achieve greatness only at the cost of profound dedication and immense sacrifice. And of course part of what we value about these cases is the product, the art or other creative object that's produced by this dedication and sacrifice. But I think we also value the process itself. There's something admirable about these figures, about their dedication, about their wholehearted and implacable commitment to their ends. Someone who can dash off a composition or splash some colors on a canvas without effort, without sacrifice, without dedication may still produce admirable work, indeed their work may be as great as that of the resolutely committed individual: but the life they live is (arguably) less admirable.[20]

I can't defend this point fully here, but I can try to make it a bit clearer. To do so, it helps to distinguish between the *end* and the *purpose* of an activity. Most activities have beginnings, middles, and ends. They have an arc, which involves progression between these things. But the *purpose* of an activity does not always correspond to its end. The end of a song is the final note, but the purpose of producing or engaging with the song is not to reach the end. It is to enjoy or appreciate the song.

So, let's look at activities that admit this end/purpose distinction, where the purpose detaches from the end. Many *playful* activities have that structure. For example, consider games, sports, and endeavors undertaken merely for the sake of the challenges they afford. Take climbing: the end is to get to the top of the rock, but that's not the purpose. Otherwise you could just use a ladder. The purpose is to face challenges, to overcome obstacles, to solve complex puzzles of alignment and positioning.[21]

In their ordinary, casual forms these activities don't require resolute commitments—it's certainly possible to have a weak commitment to climbing or yoga or playing basketball and nonetheless perform these activities quite well. But there is a form of excellence in these activities that is partly constituted by resolute commitment. When we watch an athlete's gritty pursuit of success; when we admire the dedication and commitment manifest in their efforts; when we see someone devoting themselves wholeheartedly to an activity despite hardship and sacrifice; when we see all this, we are

[20] Here I am claiming that the resoluteness is *part* of what we value when we value these lives. Of course, other values are also at issue, and sometimes they compete. For example: someone could be resolutely committed to a morally problematic end, and then the moral badness of the end might outweigh (or cancel) the goodness of the resolute commitment. Or, the end could be valuable but the resoluteness might involve the sacrifice of morally important goods, as in the Gauguin case, and some might think that this renders the resoluteness non-admirable.

[21] I offer an analysis of these kinds of cases in Katsafanas 2013. Thi Nguyen discusses the same phenomenon—but draws different lessons from it—in Nguyen 2020.

not merely admiring the attainment of some distinct end. After all, many of these ends—climbing a rock, holding a pose, winning a game—are trivial and inconsequential. What we admire is the devotion itself, the resolute commitment to the activity's purpose. And that is just what Locke's Dictum would rule out.

Finally, there are some mixed cases—cases where our interest in the resolute commitment springs both from the commitment itself and from its effects. For an abstract example, consider justice. A person who is only lackadaisically committed to justice and who lets that commitment be overridden by sufficiently weighty considerations of convenience, personal gain, happiness, and so on is not a good candidate for someone we would think of as just.[22] Insofar as we value *being just,* part of what we value is the resolute commitment. Of course, the results of the commitment also matter. But even independently of those results it seems better to be just than unjust.

So: there are cases in which resolute commitment is necessary in order to engage in a process excellently or to manifest a certain ethical or personal stance. You can sometimes get the results of the process or stance without resolute commitments, but you can't get the good of the process or stance itself.

These activities are governed by values that are contestable both in terms of their status as values and in terms of their weights or lexical orderings. So, on Locke's Dictum, they don't seem defensible targets of resolute commitment. And yet without resolute commitment they vanish.

6. Disentangling Degree of Commitment and Degree of Justification

I've argued that resolute commitments are valuable—they are implicated in some of the activities, processes, and standings that we admire. If Locke's Dictum rules out all resolute commitments, it would rule out these admirable forms of life. So we are left with two options. First, we could simply accept the idea that certain valuable features of human life depend on unreasonable commitments. Second, we could attempt to reformulate Locke's Dictum so that resolute commitments can qualify as reasonable or rational despite lacking decisive justification.

Although I'm not opposed to the first strategy, let's consider the second strategy. Here's the most obvious way of carrying it out: rather than accepting the idea that resolute commitments require an extremely high degree of justification, we say that resolute commitments demand *adequate* justification. And we then specify what count as adequate justification in a way that makes it attainable. In other words: we say that you're permitted to go all-in on a commitment once you possess adequate-but-not-decisive justification for that commitment.

So the structure is something like this. Imagine a case where it's reasonable to accept either *A* or *not-A*. If you accept *A*, it's reasonable to form a resolute commitment to *A*. If you accept *not-A*, it's reasonable to form a resolute commitment to *not-A*.

[22] There are ways of being resolutely committed that would involve dogmatism, closed-mindedness, and the inability to see competing considerations. Imagine a person who fixates on justice alone, and simply *ignores* all other relevant values and normative claims (mercy, compassion, and so on). This person might be described as resolutely committed to justice, but seems flawed. In Katsafanas 2023a, I distinguish between forms of commitment that are sustained through closed-minded, dogmatic stances and forms of commitment that, though resolute, are open-minded and inquisitive.

Or, alternatively, imagine a case where it's reasonable to prioritize either *A over B* or *B over A*. If you accept that A should be prioritized over B, it's reasonable to form a resolute commitment to prioritizing A over B. If you accept that B should be prioritized over A, it's reasonable to form a resolute commitment to prioritizing B over A.

There are examples where this seems right: I've discussed the case of loving relationships, ethical views, and artistic projects above, but consider also careers. Imagine a college student who is torn between aiming for a career as a professor and a doctor. Both are reasonable careers. To succeed in either, you typically need a reasonably strong, long-term commitment. If the student decides to aim for a career as a professor, it's reasonable for her to form a resolute commitment to this. The fact that it also would have been reasonable to form a resolute commitment to being a doctor shouldn't undermine or etiolate the actual commitment.

But notice that the most easily acceptable cases have some features that distinguish them from others: (1) they involve no cost to others (e.g. no one is harmed by your choosing one career over the other); (2) they involve no cost to yourself (e.g. you will do equally well regardless of which career you choose); (3) there is no reasonable perspective or set of assumptions from which resolute commitment to either option looks bad or mistaken.

We can distinguish this from another kind of case. Let's say that it's reasonable to accept either that animal suffering matters or that it doesn't. Then, depending on whether this claim is true, there will be costs (to the animals, obviously, but also to yourself insofar as you're committing a moral wrong) and there is a perspective from which resolute commitment looks mistaken. In a case like that, we might worry. Perhaps commitment should be hedged or weakened.

To the extent that any of those three features change, pressures begin to emerge. Suppose forming a resolute commitment involves costs to others; or involves costs to yourself; or there are competing reasonable perspectives from which going all-in on the commitment looks mistaken. Then attenuation looks more appealing. And the more so, to the extent that the degrees of these features increase. For example, the Gauguin case involves these features, and that's part of what makes it difficult. And, to circle back to the beginning: the fanatical and extremist cases involve all three of these features to much higher degrees. Again, that's what makes them so troubling.

Where, then, does this leave us? Locke's Dictum is an attempt to diagnose a problem with fanaticism by focusing on the relationship between degree of justification and degree of commitment. I've suggested that it fails: when coupled with a healthy degree of skepticism, Locke's Dictum would rule out not just fanaticism but all forms of resolute commitment. It's true that fanatical and extremist commitments are resolute; but so, too, are many valuable forms of commitment. So Locke's Dictum is mistaken, both as a general constraint on social and ethical life and as a diagnosis of fanaticism.[23]

I will close by mentioning a connection between this analysis of commitment and recent work on *faith*. So far, I've argued that Locke's Dictum should be rejected: degree of commitment should not track degree of justification. A similar point arises in certain discussions of faith.[24] Lara Buchak's influential work on faith is a good example. Buchak puts her view this way: "a proposition is only a

[23] For an alternative account of fanaticism, which avoids these problems, see Katsafanas 2023a and Katsafanas 2023b.
[24] See for example Alston 1996, Bishop 2007, Buchak 2017b, Howard-Snyder 2013, and Rettler 2018.

potential object of faith if the individual has a pro-attitude towards p and if the evidence is not enough to guarantee the truth of the proposition" (Buchak 2017a: 58). So, the person who accepts a proposition as an object of faith does not have decisive justification for believing that proposition—as Elizabeth Jackson puts it, faith has a weaker epistemic component than belief (Jackson 2021: 41). If this is right, a person could have faith that p without having sufficient justification for believing that p. If we take faith as a type of commitment, then faith violates Locke's Dictum: I can have faith that p while thinking that I lack justification for believing p.

However, there is a crucial difference between faith—at least as Buchak analyzes it—and commitment more generally. What's distinctive of faith (according to Buchak) is that it requires one to stop looking for evidence: "faith that p requires that one commit to these acts without looking for further evidence in the matter of p—at least, without looking for evidence for the sole purpose of deciding what to do —and that one maintain one's commitment even in the face of new counterevidence. Thus, faith is a matter of stopping one's search for evidence and taking action" (Buchak 2017a: 58). Commitment can be open-ended and inquisitive in a way that Buchakian faith is not. Commitment is compatible with looking for counterevidence, even very actively. But commitments need not be (and, I think, typically are not) updated whenever new evidence is encountered.[25] This is why Locke wants to defend his Dictum: it is precisely because degree of commitment often diverges from degree of justification that Locke thinks a norm regulating this relation is needed. So the feature that Buchak attributes to faith—its imperviousness to certain kinds of counterevidence—is, in fact, a feature of commitments quite generally. Thus, these reflections on Locke's Dictum help us to see that a feature that might seem peculiar to faith is in fact quite widespread.

7. Conclusion

Locke's Dictum captures something very tempting and intuitive in our thinking about the relationship between commitment and subjective justification. It tells us that these should be proportional: if your degree of justification for some commitment decreases, so too should the strength of the commitment.

Our thoughts about people of integrity and fanatics initially seem to support Locke's Dictum. However, I've suggested that a problem arises. A healthy degree of skepticism entails that our commitments always lack full-fledged justification; there will always be reasonable objections to these commitments. So Locke's Dictum combined with a healthy degree of skepticism entails that our commitments should always be attenuated. Or, put differently: Locke's Dictum entails that we ought never to have resolute commitments.

At first, this can seem salutary. After all, fanatics, enthusiasts, and extremists violate Locke's Dictum, and part of what's wrong with these figures is that they manifest extremely robust yet highly questionable commitments. So there is a temptation to think that Locke's Dictum is an appropriate normative standard. However, I've suggested that if we abided by Locke's Dictum, certain valuable features of human life would disappear. For many valuable features of ethical and social life depend on resolute commitments.

[25] See also note 17 above.

So we either need to accept that valuable features of human life depend on unreasonable commitments or we need to disentangle degree of commitment and degree of justification, positing a more complex relationship between them. Pursuing the latter option, we could say that once justification reaches a certain threshold, it's reasonable to form a resolute commitment. But this picture is less tempting when resolute commitments involve costs to oneself, costs to others, or look mistaken from other reasonable perspectives.

Acknowledgements

For extremely helpful comments on earlier drafts of the paper, thanks to Madison Favela, Josh DiPaolo, Quassim Cassam, Susanna Siegel, Heather Battaly, Tracy Llanera, Robert Simpson, Michaela McSweeney, Aaron Garrett, Juliet Floyd, Daniel Dahlstrom, Pol Pardini, Walter Hopp, Daniel Star, Jeremy Fix, Alex Prescott-Couch, Rachel Achs, Sigrun Svavarsdottir, Erin Kelly, Avner Baz, Dilip Ninan, Lionel McPherson, Christiana Olfert, Ian Blaustein, anonymous reviewers for this journal, and audiences at CSU Fullerton, Tufts University, University of Essex, and Oxford University. This publication was made possible through the support of Grant 62871 from the John Templeton Foundation. The opinions expressed in this publication are those of the author and do not necessarily reflect the views of the John Templeton Foundation.

References

Alston, William (1996). Belief, Acceptance, and Religious Faith. In J. Jordan and D. Howard-Snyder (eds.), *Faith, Freedom, and Rationality* (3-27). Rowman and Littlefield

Andreou, Chrisoula (2022). *Commitment and Resoluteness in Rational Choice.* Cambridge University Press.

Anstey, Peter R. (2019). Locke, the Quakers and Enthusiasm. *Intellectual History Review* 29 (2): 199–217.

Battaly, Heather (2023). Can Fanaticism be a Liberatory Virtue? *Synthese* 201(6): 1-27

Bishop, John (2007). *Believing by Faith*. Clarendon Press.

Bratman, Michael (2012). Time, Rationality, and Self-Governance. *Philosophical Issues* 22: 73–88.

Bratman, Michael (2018). *Planning, Time, and Self-Governance*. Oxford University Press.

Buchak, Lara (2017a). Reason and Faith. In Frederick D. Aquino (ed), *The Oxford Handbook of Epistemology of Theology* (46-63). Oxford University Press.

Buchak, Lara (2017b). Faith and Steadfastness in the Face of Counter-evidence. *The International Journal for Philosophy of Religion* 81(1): 113–133

Calhoun, Cheshire (1995). Standing for Something. *Journal of Philosophy* 42: 235–260.

Calhoun, Cheshire (2018). *Doing Valuable Time: The Present, the Future, and Meaningful Living.* Oxford University Press.

Cassam, Quassim (2022). *Extremism: A Philosophical Analysis.* Routledge.

Casson, Douglas (2024). Locke on Fanaticism. In Paul Katsafanas (ed.), *The History and Philosophy of Fanaticism* (83-96). Routledge.

Chang, Ruth (2013). Commitments, Reasons, and the Will. *Oxford Studies in Metaethics* 8: 74-113.

DiPaolo, Joshua (2020). The Fragile Epistemology of Fanaticism. In Michael Klenk (ed.), *Higher-Order Evidence and Moral Epistemology* (217-235). Routledge.

Frankfurt, Harry (2004). *The Reasons of Love*. Princeton University Press.

Halfon, Mark (1989). *Integrity: A Philosophical Inquiry*. Temple University Press.

Hébert, M. R. (2002). Integrity, Identity, and Fanaticism. *Contemporary Philosophy* 24: 25-29.
Holton, Richard (2009). *Willing, Wanting, Waiting*. Oxford University Press.
Howard, Christopher (2019). Fitting Love and Reasons for Loving. In Mark Timmons (ed.) *Oxford Studies in Normative Ethics, Volume 9* (116-137). Oxford University Press.
Howard-Snyder, Daniel (2013). Propositional Faith: What It Is and What It Is Not. *American Philosophical Quarterly* 50(4): 357–372.
Jackson, Elizabeth (2021). Belief, Faith, and Hope: On the Rationality of Long-Term Commitment. *Mind* 130: 35-57.
Katsafanas, Paul (2013). *Agency and the Foundations of Ethics*. Oxford University Press.
Katsafanas, Paul (2023a). *Philosophy of Devotion: The Longing for Invulnerable Ideals*. Oxford University Press.
Katsafanas, Paul (2023b). What's So Bad About Fanaticism? *Synthese* 203: 1-18.
Kolodny, Niko (2003). Love as Valuing a Relationship. *The Philosophical Review* 112 (2): 135–89.
Korsgaard, Christine (1996). *The Sources of Normativity*. Cambridge University Press.
Locke, Dustin (2015). "Practical Certainty", *Philosophy and Phenomenological Research* 90(1): 72–95
Locke, John (1978). *The Correspondence of John Locke*. Edited by E.S. de Beer. 8 vols. Oxford University Press.
Locke, John (2000). *An Essay Concerning Human Understanding*. Batoche Books.
Luther, Martin (1525). Against the Murderous, Thieving Hordes of Peasants. In Helmut T. Lehmann, Robert C. Schultz, and Joel W. Lundeen (eds.), *Luther's Works, Volume 46* (47-55). Fortress Press.
Martin, Adrienne (2013). *How We Hope: A Moral Psychology*. Princeton University Press.
McFall, Lynne (1992). Review of *Integrity: A Philosophical Inquiry* by Mark S. Halfon. *The Philosophical Review* 101(2): 463-65.
Morton, Jennifer M. and Sarah K. Paul (2019. Grit. *Ethics* 129(2): 175–203.
Nguyen, Thi (2020). *Games: Agency as Art*. Oxford University Press.
Protasi, Sara (2016). Loving People for Who They Are (Even When They Don't Love You Back). *European Journal of Philosophy* 24 (1): 214–34.
Rettler, Bradley (2018). Analysis of Faith. *Philosophy Compass* 13(9): 1–10
Rioux, Catherine (2022). A Higher-Order Approach to Diachronic Continence. *Thought* 11(1): 51-58.
Scheffler, Samuel (2013). *Death and the Afterlife*. Oxford University Press.
Scherkoske, Greg (2013). *Integrity and the Virtues of Reason: Leading a Convincing Life*. Cambridge University Press.
Svavarsdóttir, Sigrún (2015). Coherence of Attitudes, Integration of the Self, and Personal Identity. In David Shoemaker (ed.), *Oxford Studies in Agency and Responsibility, Volume 3*. Oxford University Press.
Svavarsdóttir, Sigrún (2022). Integrity, Truth, and Value. In Tamar Shapiro, Sharon Street, and Kyla Ebels-Duggan (eds.) *Normativity and Agency: Themes from the Work of Christine Korsgaard* (147-174). Oxford University Press.
Taylor, Gabrielle (1981). Integrity. *Aristotelian Society Supplementary Volume* 55 (1): 143–176.
Velleman, J. David (1999). Love as a Moral Emotion. *Ethics* 109: 338–74.
Voltaire (1962). *Philosophical Dictionary,* ed. Peter Gay. Basic Books.
Williams, Bernard (1981). Moral Luck. In *Moral Luck* (20-39). Cambridge University Press.
Wong, David (2006). *Natural Moralities: A Defense of Pluralistic Relativism*. Oxford University Press.