

Forthcoming in *The Philosophical Review*. Penultimate draft.

## **A Faithful Response to Disagreement\***

Lara Buchak

**ABSTRACT:** In the peer disagreement debate, three intuitively attractive claims seem to conflict: there is disagreement among peers on many important matters; peer disagreement is a serious challenge to one's own opinion; and yet one should be able to maintain one's opinion on important matters. I show that contrary to initial appearances, we can accept all three of these claims. Disagreement significantly shifts the balance of the evidence; but with respect to certain kinds of claims, one should nonetheless retain one's beliefs. And one should retain them even though these beliefs would not be supported by the new total evidence if one didn't already hold them.

**KEYWORDS:** disagreement, faith, belief, credence, counterevidence

### **1. Introduction**

Sometimes you encounter someone who is your intellectual equal and has roughly the same evidence you have, but who has come to a very different conclusion on the basis of that evidence. This is the setting for the peer disagreement debate. The question is how this encounter should change your opinion. Three intuitively attractive claims seem to conflict: there is disagreement among peers on many important matters; peer disagreement is a serious challenge to one's own opinion; and yet one should be able to maintain one's opinion on important matters. However, these claims only conflict if we accept an assumption about how one's opinion should be related to the evidence one has at a time. Although this assumption seems so obvious that it goes unquestioned in the debate, I will show that it can be rejected,

---

\* This paper benefited greatly from comments by Dan Speak, Liz Jackson, and Sven Neth; from conversations with audiences at the Midwest Society of Christian Philosophers 2015, the Templeton Faith Summer Seminar at the University of Missouri, the Pacific APA 2016, the Mountain-Pacific Society of Christian Philosophers 2017, the Biola Center for Christian Thought, and Syracuse University; and from discussion in seminars at Rutgers, Ryerson, and Princeton. It was also made possible through the support of a grant 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).

and rejected precisely in the case of peer disagreement about important matters. This opens up a promising response to the problem of peer disagreement. Peer disagreement does give you evidence against the claim which is the object of disagreement, but you should nonetheless retain your initial belief.

## 2. Peer Disagreement

In the partial-belief version of peer disagreement, an individual examines a body of evidence and on its basis forms a credence in some claim ‘X’. She then encounters an *epistemic peer*—an individual who has roughly the same cognitive abilities and who has examined roughly the same evidence—who forms a different credence.<sup>1</sup> The question at issue is how she should update her credence in response to the disagreement. In the full-belief version, the individual forms a belief that ‘X’ (or that ‘not-X’), and meets a peer who believes the opposite; and the question at issue is what she should believe after the encounter.

Three claims about this problem have initial appeal. The first is that disagreement with epistemic peers is widespread, particularly with respect to important matters like morality, religion, and politics.<sup>2</sup> Indeed, this is why peer disagreement is taken to be a topic of philosophical interest in the first place. Call the

---

<sup>1</sup> Here is a canonical statement of epistemic peerhood: “Let’s say that people are *epistemic peers* when they are roughly equal with respect to intelligence, reasoning powers, background information, etc. When people have had a full discussion of a topic and have not withheld relevant information, we say that they have *shared their evidence* about that topic”(Feldman 2007). I prefer a more general characterization, in particular one in which the peer’s evidence isn’t the same but is ‘just as good’. (See King (2012) for an argument that peers rarely have the same evidence.) An example of the kind of characterization I favor is that of Pittard (2020), who defines two people as epistemic peers with respect to *p* “just in case their epistemic credentials with respect to *p* are equally strong, where strength should be understood as taking into account *all* of the dimensions of epistemic evaluation that bear on the reliability of one’s judgment, including facts about the quality and quantity of one’s evidence as well as facts about one’s ability to respond to that evidence in a rational and unbiased manner”(26).

<sup>2</sup> This is either explicitly stated or presupposed by van Inwagen (1996), Rosen (2001), Kelly (2005), Feldman (2007), Christensen (2009), Kelly (2010), Kornblith (2010), Lackey (2010a), Lackey (2010b), and van Inwagen (2010). Earlier discussions grew out of issues surrounding religious pluralism, and concerned the challenge of disagreement to religious belief in particular: a classic paper is van Inwagen (1996). Some of the recent literature has similarly narrowed the focus to one particular domain: Kraft (2007), Bogardus (2013), Pittard (2014), Bergmann (2017), Dormandy (2018), Loughheed (2018b), Pittard (2020), Choo (forthcoming), and Turnbull (forthcoming), among others, focus on religious disagreement; McGrath (2008), Vavova (2014), Fritz and McPherson (2019) focus on moral disagreement; Hallsson (2019) focuses on political disagreement; Fleisher (2018) focuses on scientific disagreement; and Goldberg (2013b), Barnett (2019), and Palmira (2019) focus on philosophical disagreement.

target claims—the claims about important matters in these domains that are putatively the subject of peer disagreement—the DISPUTED CLAIMS. It is not difficult to come up with examples of DISPUTED CLAIMS: that consequentialism is correct, that abortion is morally impermissible, that God exists, that some particular religion is broadly speaking true, that human activity is causing drastic climate change. Thus:

FREQUENCY: There is widespread disagreement among peers about the DISPUTED CLAIMS.

The second appealing claim is that peer disagreement presents a serious challenge to one's own opinion (where here I use 'opinion' to mean either credence or belief, depending on which attitude is at issue):

HUMILITY: When an epistemic peer forms a very different opinion about a claim than I do, I ought to drastically alter my opinion.

Many philosophers—those known as *conciliationists* or *conformists*—continue to endorse this claim after careful argument. Even those who ultimately reject this claim take its rejection to be something in need of argument, because of its appeal. For example, Thomas Kelly states that “it is natural to suppose that persistent disagreement among epistemic peers should undermine the confidence of each of the parties in his or her own view.”<sup>3</sup> A typical reason, though not necessarily the only reason, to endorse this claim is because one's opinion and that of one's peer are symmetrical in some way and thus one must give one's peer's opinion equal weight as one's own, or at least significant weight.<sup>4</sup>

The third appealing claim is that we would lose something valuable if we could not stick to our guns on topics of central importance:<sup>5</sup>

TENACITY: I ought to maintain my opinion about the DISPUTED CLAIMS.

---

<sup>3</sup> Kelly (2005: 3). Kelly also cites several philosophers from the past who voice this intuition, including Sextus Empiricus, Montaigne, and Sidgwick.

<sup>4</sup> That considerations of symmetry underlie HUMILITY is endorsed by Feldman (2007), Kelly (2005), and Christensen (2007), among others.

<sup>5</sup> Explicit statements of this claim are found in van Inwagen (1996), Kelly (2005), Pettit (2006), Christensen (2007), Elga (2007), Christensen (2009), Vavova (2014), Fritz (2018), and Fritz and McPherson (2019).

Unfortunately, these three claims form an inconsistent triad. If peers disagree about the disputed claims, and if I must alter my opinion in response to peer disagreement, then I cannot maintain my opinion about these claims.

One way to sort the peer disagreement literature is in terms of which of these three claims an author denies. Two groups of philosophers deny FREQUENCY. Some deny FREQUENCY on the grounds that when it comes to important matters, one should rarely count those who disagree as being equally good reasoners. This could be because one cannot evaluate an apparent peer's track record independently of the truth of the many other things you disagree about,<sup>6</sup> or because one cannot evaluate his epistemic qualifications independently of the truth of the claim at issue.<sup>7</sup> Alternatively, some deny FREQUENCY on the grounds that there is often an important evidential asymmetry between apparent peers.<sup>8</sup>

Other philosophers deny HUMILITY. Many of these philosophers concentrate on an asymmetry between one and one's peer, from the point of view of each: an incommunicable insight,<sup>9</sup> a non-doxastic appearing-true,<sup>10</sup> one's judgment of the probative force of the evidence,<sup>11</sup> one's own perspective on the shared evidence and reasons,<sup>12</sup> or one's reasonable view of the evidential norms.<sup>13</sup> (Note that although these views are most naturally read as denying HUMILITY, they could also be read as denying

---

<sup>6</sup> Pettit (2005), Elga (2007).

<sup>7</sup> Pittard (2014).

<sup>8</sup> Fumerton (2010) notes that when he is confident of his philosophical or political views, he is also confident that he can eventually bring other around people he thinks of as reasonable, if he leads them through his progression of thought. Goldman (2010) notes that people who disagree often have different evidence about which norms of reasoning are correct; so if 'norm evidence' is included in the characterization of evidence, then there are fewer peers than we thought.

<sup>9</sup> Van Inwagen (1996), van Inwagen (2010) (though the latter does not fully endorse this view), Loughheed (2018b). See also Pittard's (2020) argument that a *rational* insight can be a symmetry-breaker.

<sup>10</sup> Rosen (2001). See also Bergmann (2017).

<sup>11</sup> Kelly (2005).

<sup>12</sup> Conee (2010). See also Fritz and McPherson (2019), who suggest, but don't fully endorse, the idea that it is epistemically virtuous to maintain one's belief because it was formed autonomously or facilitates one's understanding of the matter at hand.

<sup>13</sup> Wedgwood (2010). This is also the position of Goldman (2010), if 'norm evidence' is not included in the characterization of evidence. See footnote 9 in this paper. Elgin (2018) also notes several of the asymmetries mentioned above.

FREQUENCY: to the extent that our conception of what evidence consists in includes any of these factors, one's apparent peer will share less evidence.) Others deny HUMILITY by holding that there is an asymmetry that could justify one peer without also justifying the other, so that there is no general norm about whether one should stick to one's view or revise in the face of disagreement. On Thomas Kelly's Total Evidence View,<sup>14</sup> one's total evidence, including the evidence provided by the presence of the disagreeing peer, determines what one's ultimate belief state should be: if one's initial evidence strongly supported one's initial conclusion, then one may not need to revise very much, if at all. On Jennifer Lackey's Justificationist View,<sup>15</sup> one's initial degree of justified confidence determines whether disagreement is significant, with the result that one may not need to revise one's beliefs at all.

Finally, some philosophers deny TENACITY. Some of these philosophers hold that this is a disappointing conclusion—Richard Feldman calls it “distressing” and Hilary Kornbluth accepts it “very reluctantly”—but it is where we are pushed by the truth of FREQUENCY and HUMILITY.<sup>16</sup> Others try to draw a positive moral from this conclusion, for example that the beliefs of others can serve as “epistemic checks” on one's own<sup>17</sup> or can reveal that there is no reliable route to truth in the relevant domain.<sup>18</sup> More recently, some have argued that we cannot maintain our opinions in the face of disagreement, but that we can sometimes maintain a different attitude: speculating,<sup>19</sup> suspecting,<sup>20</sup> endorsing,<sup>21</sup> or hypothesizing.<sup>22</sup>

Notably, the same types of arguments are offered by those who take ‘opinion’ to be belief, disbelief, or agnosticism, and those who take ‘opinion’ to be various levels of credence (‘partial’ belief).

I will not evaluate any of the positions mentioned. It is certainly possible to deny one of the above claims, initial appeal notwithstanding. My point will be that there is another response: we can accept a

---

<sup>14</sup> Kelly (2007, ) 2010, (2013).

<sup>15</sup> Lackey (2010a, 2010b).

<sup>16</sup> Feldman (2007), Kornblith (2010).

<sup>17</sup> Christensen (2007). See also Elgin (2018).

<sup>18</sup> Goldberg (2009, 2013a, 2014).

<sup>19</sup> Goldberg (2013b), Barnett (2019).

<sup>20</sup> Carter (2018).

<sup>21</sup> Fleisher (2018).

<sup>22</sup> Palmira (2019).

version of all three claims, by accepting HUMILITY when it comes to the evidence for a claim, and accepting TENACITY when it comes to believing the claim. That is, we can accept:

FREQUENCY: There is widespread disagreement among peers about the DISPUTED CLAIMS.

HUMILITY (EVIDENCE): When an epistemic peer forms a very different opinion about a claim than I do, the balance of my evidence for that claim changes significantly.

TENACITY (BELIEF): I ought to maintain my belief about the DISPUTED CLAIMS.

Note that these claims are only incompatible if we accept:

SYNCHRONICITY: My belief after I encounter peer disagreement should depend only on the balance of my evidence after I encounter peer disagreement.

We will deny SYNCHRONICITY by holding that while in general one's belief should depend on the evidence, one's belief needn't *at every time* depend on the balance of the evidence *at that time*. In particular, there are circumstances in which one's belief ought to swing free from the evidence one has at some particular times in favor of past evidence—circumstances that are particularly apt to obtain with respect to peer disagreement about important matters. Seeing why this is the case will point us to a promising and unexplored alternative in the peer disagreement debate. The argument here also has two other important upshots. First, it can serve as a defense for a limited kind of conservatism in belief (conservatism when it comes to beliefs about important matters). And second, since 'the balance of the evidence' is reflected in one's *credence*, it will show that we can hold one view about disagreement when it comes to credence and another when it comes to belief.<sup>23</sup>

### 3. Belief and Convictions

In this paper I will present a very simple formal model that makes sense of FREQUENCY, HUMILITY (EVIDENCE), and TENACITY (BELIEF) by showing when and why we should reject

---

<sup>23</sup> Jackson (2019) points out that which attitude type is under discussion in the debate about peer disagreement makes a difference to how the debate is constrained.

SYNCHRONICITY. Of course, the aim here isn't formal modeling itself. The point of the formal model is to introduce a (perhaps unrealistic) degree of precision in order to draw our attention to a particular phenomenon in the messy real world. So the important feature of the model will be *why* we should sometimes reject SYNCHRONICITY.

As is standard in formal epistemology, the balance of the evidence will be represented by a subjective probability (*credence*) function, and a rational credence function will (at least) obey the probability calculus and be updated via conditionalization. (We are not here concerned with what mental feature, exactly, credence represents; instead, we care that credence reflects the balance of the evidence.<sup>24</sup>) We will adopt a view of rational *belief* according to which belief balances the fulfillment of two cognitive goals: believing truth and avoiding error.<sup>25</sup> In particular, we will adopt Isaac Levi's framework that models accepting a proposition (in our terminology: believing a proposition) as an action whose utility payoff depends on the truth of the proposition believed.<sup>26</sup> In this framework, one's 'cognitive options' are distinguished by the logically strongest proposition one accepts. (Abstaining from accepting any proposition—being agnostic—is regarded as believing a logically weak proposition that is guaranteed to be true, e.g., '*X or not-X*'.) One gets a higher utility payoff if one selects a true cognitive option than a false one, and if one selects a true cognitive option, the utility payoff of that option is higher the logically stronger it is.

For example, consider an individual trying to decide what to accept about which of three candidates will win an election. It would be best to accept of the winning candidate that she will win, next best to accept that she or some rival will win, next best to be agnostic, and worst to accept something false about who will win (even worse the weaker it is). So, if candidate A wins, then the payoffs are, in order:

---

<sup>24</sup> See Joyce (2005) on how different features of credence reflect different properties of the evidence.

<sup>25</sup> See James (1896).

<sup>26</sup> Levi (1967a). See also Levi (1967b). Levi embeds this framework in a complex view about knowledge (Levi (1980, 1991)); here I just make use of the elements of his framework described in this section, and make no commitments about the surrounding picture. Fleisher (2018) makes use of a similar model to explain why a scientist should *endorse* her favored view even when her credence in that view is lower than in its negation (and even when she receives counterevidence to her view, the situation I focus on here and in Buchak (2017)). On his view, endorsement and its governing norms are different from belief and its governing norms in that endorsement is specific to a research domain, fragmented, compartmentalized, and does not aim directly at truth; and that it might be better to endorse a false theory than a true one. But endorsement shares some features with the attitude I characterize as faith, most importantly that it involves resilient commitment, and is more resilient than belief.

$$\begin{aligned}
 &u(\text{Believe 'A will win'}) \\
 &u(\text{Believe 'A or B will win'}) \quad u(\text{Believe 'A or C will win'}) \\
 &u(\text{Believe 'A or B or C will win'}) \\
 &u(\text{Believe 'B will win'}) \quad u(\text{Believe 'C will win'}) \\
 &u(\text{Believe 'B or C will win'})
 \end{aligned}$$

And similarly for the payoffs if B wins or if C wins. So there are three possible states of the world in which each proposition is evaluated. I will talk about the simpler case in which there is only one proposition at issue, so there are three cognitive options: believe 'X', believe 'not-X', or remain agnostic.

Levi himself gives a story about how we might set these utility values, which includes a parameter that in the single-proposition case measures how close the payoff of agnosticism is to that of believing a falsehood. Thus, we can without loss of generality set the utility of believing 'X' when  $X$  holds, and of believing 'not-X' when not- $X$  holds, to 1; set the utility of believing 'X' when not- $X$  holds, and of believing 'not-X' when  $X$  holds, to 0; and set the utility of agnosticism to some middle value  $M$ , where  $0.5 \leq M < 1$ . Accepting a logically strong proposition is therefore risky. Relative to being agnostic, it has a higher payoff if the proposition is true and a lower payoff if the proposition is false.

Given a subject's credences, we can evaluate each cognitive option using decision theory. In standard decision theory, the value of each option is its expected utility relative to the subject's credences.<sup>27</sup> So, for example, if we set  $M = 0.75$ , then we require a credence in 'X' of greater than 0.75 to believe 'X', and similarly for 'not-X'. Note that this is *not* the 'threshold' view (according to which belief is credence above a threshold), because when multiple propositions are at issue, the minimum credence required to believe 'X' will depend on both the propositions involved and on what else the subject accepts.

This picture makes belief subject to the same standards as action, and thus allows us to use decision theory to say how one's beliefs should relate to one's credences. If we think of credence as an estimate of truth-value, the rationale is clear: one's goals with respect to belief are to believe truths and avoid

---

<sup>27</sup> Levi himself eventually rejects standard decision theory because he rejects the idea that a rational credal state must be represented by a single credence function (Levi 1980), but since I use this model only to illustrate why we can reject SYNCHRONICITY, this does not matter for our purposes. Note also Levi's rule for ties (Levi 1967a: 84), in which agnosticism is privileged over belief.

believing falsehoods, and so one uses one's estimate of truth-value to pick the cognitive option that is does best with respect to these goals. This picture also relates one's beliefs and one's evidence, in a mediated way: one's credence is supposed to reflect the balance of the evidence, and so one picks the cognitive option that does best with respect to the two goals, given one's evidence.<sup>28</sup>

Does this picture require that belief is 'up to us' or a matter of the will? Not necessarily. It is standard to distinguish between an *evaluative* decision theory—a theory that gives the conditions under which someone is acting rationally—and an *action-guiding* decision theory—a theory that gives followable advice about how one should act. Here we use the model evaluatively: if a belief maximizes a subject's cognitive utility, then it is the belief she should have.<sup>29</sup> But that is compatible with belief not being under one's control at all, in which case the picture here will tell us how to evaluate various automatic mechanisms that produce belief—indeed, as I will show, it can explain how a certain seemingly irrational mechanism (conservatism about belief) is in fact rational.

#### 4. Updating on Evidence without Believing according to it

In a series of articles, I have argued that there are circumstances in which it is rational, at the current time, to *commit* to taking a risk on *X*, even if one might in the future receive counterevidence that will make one no longer want to take the risk; and furthermore, it is also rational to *follow through* on that commitment when the time comes. (Here, taking a risk on *X* means choosing an act that will yield high utility if *X* and low utility if not-*X* rather than an alternative course of action that will yield intermediate utility either way.) In other words, it can be rational to choose the plan of acting on *X* in the future rather than the plan of basing one's future action on future evidence, and to follow through on that plan when the time comes, even if one would not have adopted the plan anew when the time comes.<sup>30</sup> I hold that

---

<sup>28</sup> If one is worried that this picture mediates belief through something subjective, one could instead hold that there is an objectively correct credence function, that one's subjective credence function ought to coincide with it, and that one's beliefs should maximize cognitive utility with respect to the *correct* credences. This suggestion makes it possible to capture Maria Lasonen-Aarnio's (2014) view that one's belief should depend on *the evidence* itself, not on one's assessment thereof (where 'evidence itself' would be represented by the objective credence function and 'assessment thereof' would be represented by the subjective credence function). However, the suggestion makes it difficult to apply this picture to cases of 'pure' peer disagreement, where two peers disagree on what credence is warranted by a single body of evidence.

<sup>29</sup> This is also Levi's view: see, e.g., Levi (1967a: 21).

<sup>30</sup> See Buchak (2010, 2012, 2014, 2017).

committing, and maintaining one’s commitment, to act on a claim constitutes *having faith* in that claim, but for our purposes it does not matter if I accurately characterize faith—what matters is the conditions under which it is rational to ignore counterevidence when acting.

Certain of these circumstances apply to belief, analyzed according to the theory in the previous section. Again, assume that one is concerned with believing the truth about a single proposition  $X$ . In this case, the relevant risk is *believing* ‘ $X$ ’ or ‘not- $X$ ’ (Figure 1). Thus, the result will be about when it is rational to commit to believe ‘ $X$ ’ rather than to plan to update one’s beliefs in response to new evidence.

	X	Not-X
<b>Believe ‘X’</b>	<b>1</b>	<b>0</b>
Be agnostic	M	M
Believe ‘not-X’	0	1

FIGURE 1. Payoffs of Various Belief States,  $0.5 \leq M < 1$

We will assume that at some initial time  $t_1$ , one’s credence in ‘ $X$ ’ is high enough to make “believe ‘ $X$ ’” the option with the highest utility.<sup>31</sup> Now, assume that one will receive a piece of evidence, either  $E$  or not- $E$ , where  $E$  (‘positive evidence’) raises one’s credence in ‘ $X$ ’ and not- $E$  (‘negative evidence’) lowers it, after which (time  $t_2$ ) we can again evaluate which option has the highest utility. Clearly, if one receives  $E$ , then “believe ‘ $X$ ’” still has the highest utility.<sup>32</sup> We will restrict our discussion to the scenario in which if one receives not- $E$ , then “be agnostic” has the highest utility (Figure 2).<sup>33</sup>

	X & not-E	Not-X & not-E
Believe ‘X’	1	0
<b>Be agnostic</b>	<b>M</b>	<b>M</b>
Believe ‘not-X’	0	1

FIGURE 2. Payoffs after Negative Evidence

We want to think about how an individual should *plan* to believe over time in this scenario. The choice at  $t_1$  between, on the one hand, committing and, on the other, planning to re-evaluate after additional

<sup>31</sup> For an expected utility maximizer, this means that  $p(X)(1) + p(\text{not-}X)(0) > M$ , i.e.,  $p(X) > M$ .

<sup>32</sup> This assumes that increasing the probability of the better outcome does not make a gamble worse, a fairly innocuous assumption shared by both expected utility maximization and the theory mentioned in footnote [36].

<sup>33</sup> For an expected utility maximizer, this means that  $p(X | \text{not-}E) \leq M$  and  $p(\text{not-}X | \text{not-}E) \leq M$ , i.e.  $1 - M \leq p(X | \text{not-}E) \leq M$ . We are assuming that if utilities are tied, agnosticism is preferred to belief.

evidence is represented in Figure 3, since ‘re-evaluation’ becomes agnosticism if one receives negative evidence.

	X & E ‘correctly leading positive evidence’	X & not-E ‘misleading negative evidence’	Not-X & E ‘misleading positive evidence’	Not-X & not-E ‘correctly leading negative evidence’
Commit to believe ‘X’	1	1	0	0
Believe ‘X’ and re-evaluate in response to negative evidence	1	M	0	M

FIGURE 3. Payoffs of Committing to Believe ‘X’ versus Re-evaluating

I use ‘positive’ and ‘negative’ evidence to refer to evidence that tells in favor of and against  $X$ ; and I use ‘correctly leading’ and ‘misleading’ evidence to refer to evidence that tells in favor of and against *the truth*, whatever that happens to be. Notice that the sense in which evidence might be misleading is that it might be misleading *about X*—I am assuming that if the agent gets evidence  $E$  then  $E$  in fact holds. For example, when a third party tells you your friend is a thief, you learn the truth about what the third party said—but if your friend is not a thief, then your credence is pulled away from the truth about *this* claim, so the evidence is misleading about *it*.

As we can see, each of the acts has a higher payoff than the other in some states. In the state in which one receives correctly leading negative evidence, it is better to re-evaluate than to commit, and it is better by utility  $M$  (the difference between being agnostic and believing a falsehood). This is the benefit of *being talked out of a bad risk on X*. In the state in which one receives misleading negative evidence, it is better to commit than to re-evaluate, and it is better by utility  $1 - M$  (the difference between believing the truth and being agnostic). This is the benefit of *not being talked out of a good risk on X*, or of *sticking with a good risk when you would otherwise be talked out of it*.

For an expected utility maximizer, the benefit of being talked out of a bad risk will never be outweighed by the benefit of not being talked out of a good risk.<sup>34</sup> The former benefit will outweigh the latter

<sup>34</sup>  $EU(\text{commit}) = p(X \& E)(1) + p(X \& \text{not-}E)(1) + p(\text{not-}X \& E)(0) + p(\text{not-}X \& \text{not-}E)(0)$

$EU(\text{believe \& re-evaluate}) = p(X \& E)(1) + p(X \& \text{not-}E)(M) + p(\text{not-}X \& E)(0) + p(\text{not-}X \& \text{not-}E)(M)$

$EU(\text{believe \& re-evaluate}) \geq EU(\text{commit})$

$\Leftrightarrow p(\text{not-}X \& \text{not-}E)(M - 0) \geq p(X \& \text{not-}E)(1 - M)$

benefit more the higher  $p(\text{not-}X \ \& \ \text{not-}E)$  is relative to  $p(X \ \& \ \text{not-}E)$ , i.e., the higher the probability of correctly leading negative evidence relative to the probability of misleading negative evidence. Holding fixed  $p(\text{not-}E)$ , the former benefit will outweigh the latter benefit the higher  $p(\text{not-}X \mid \text{not-}E)$  is, i.e., the better correlated the negative evidence is with the falsity of  $X$ .<sup>35</sup>

Thus, in the setup above, committing to believe ‘X’ cannot be better than planning to re-evaluate, no matter what the probabilities. However, there are two cases in which committing can be better, for particular probabilities.

The first case is one in which the individual does not maximize expected utility, but instead weights worse scenarios more heavily than better scenarios.<sup>36</sup> Since the existence of rational non-expected utility maximizers is controversial, I will not treat this case here.

Instead, I will concentrate on the second case: *there are epistemic benefits to committing to a true belief beyond the benefits of currently believing the truth with respect to that proposition.* I will present an extremely simple model to illustrate the basic idea.

Let  $b$  be the benefit, at  $t_1$ , of committing to believe  $X$  if  $X$  holds (assume  $b > 0$ ). Then the decision, at  $t_1$ , between committing and planning to re-evaluate is represented in Figure 4.

---


$$\Leftrightarrow p(\text{not-}E)p(\text{not-}X \mid \text{not-}E)(M) \geq p(\text{not-}E)p(X \mid \text{not-}E)(1 - M)$$

$$\Leftrightarrow p(\text{not-}X \mid \text{not-}E)(M) \geq p(X \mid \text{not-}E)(1 - M)$$

$$\Leftrightarrow (1 - p(X \mid \text{not-}E))(M) \geq p(X \mid \text{not-}E)(1 - M)$$

$$\Leftrightarrow M \geq p(X \mid \text{not-}E)$$

This final inequality holds by supposition that ‘Be agnostic’ is weakly preferred after receiving *not-E* (see previous footnote). If  $M > p(X \mid \text{not-}E)$ , i.e. if agnosticism is strictly preferred, then  $EU(\text{believe} \ \& \ \text{re-evaluate})$  is strictly higher than  $EU(\text{commit})$ .

This is a case of a more general theorem by I.J. Good (1967).

<sup>35</sup> If  $p(\text{not-}X \mid \text{not-}E) = 1 - M$ , then the benefit of being talked out of a bad risk is equal to the benefit of not being talked out of a good risk, with the former increasing relative to the latter, as  $p(\text{not-}X \mid \text{not-}E)$  increases..

<sup>36</sup> See Buchak (2013).

	X & E 'correctly leading positive evidence'	X & not-E 'misleading negative evidence'	Not-X & E 'misleading positive evidence'	Not-X & not-E 'correctly leading negative evidence'
Commit to believe 'X'	1 + b	1 + b	0	0
Believe 'X' and re-evaluate in response to negative evidence	1	M	0	M

FIGURE 4. Payoffs of Committing to Believe 'X' versus Re-evaluating, with Commitment Benefits

As before, committing to believe 'X' has the benefit of not being talked out of a good risk, to be weighed against the benefit of being talked out of a bad risk; while the latter is always higher, they approach equality as  $p(X | \text{not-E})$  approaches  $M$ . (Call the difference between these values *the benefit of acting on negative evidence*.) Now, in addition, there is a benefit to committing ( $b$ ), which obtains in all of the circumstances in which  $X$  obtains (i.e., in  $X \& E$  as well as in  $X \& \text{not-E}$ ). The positive contribution of this benefit will thus be higher when  $p(X)$  is higher, and it could in principle be high enough to outweigh the benefit of acting on negative evidence, if  $p(X)$  is high enough and  $p(X | \text{not-E})$  is high enough (without being so high as to make switching unattractive in the first place).<sup>37</sup>

Let us make this concrete by assuming that  $p(X | \text{not-E}) = 0.5$ . If  $M = 0.5$  (likely unrealistic, but a limit case), then the benefit of acting on negative evidence will be 0, so if there is any benefit at all to commitment to the truth, it will be better to commit to believe 'X' than to re-evaluate.<sup>38</sup> As  $M$  gets larger, the benefit of acting on negative evidence increases (intuitively: as the value of agnosticism gets closer to the value of a true belief, and farther from the value of a false belief, the good of being talked out of a bad risk is greater than the bad of being talked out of a good risk), so there will need to be more benefit to commitment or antecedently higher  $p(X)$  to make commitment better than re-evaluation.

<sup>37</sup> In this case,  $EU(\text{commit}) - EU(\text{believe \& re-evaluate}) = p(\text{not-E})[p(X | \text{not-E}) - M] + p(X)b$ . This is positive when  $p(X | \text{not-E}) = M$  and  $p(X)$  and  $b$  are both positive. As  $p(\text{not-E})$  increases and  $p(X | \text{not-E})$  decreases, we require a larger  $p(X)$  and/or a larger  $b$  for  $EU(\text{commit}) - EU(\text{believe \& re-evaluate})$  to remain positive. See also Buchak (2012).

<sup>38</sup> See previous footnote. Plugging in the values here, we get:  $EU(\text{commit}) - EU(\text{believe \& re-evaluate}) = p(\text{not-E})(0) + p(X)b = p(X)b$ .

How should we measure the value of commitment? I don't have a concrete proposal, but the thing to do is to ask how the various commitment benefits talked about later compare to, say, the advantage of true belief over agnosticism.

What happens at time  $t_2$ , if one chooses the option of commitment and in fact does receive negative evidence? *As long as the benefits of commitment still remain*, then the individual faces the decision problem in Figure 5.

	X & not-E Misleading negative evidence	Not-X & not-E Correctly leading negative evidence
Believe 'X' (stick with commitment)	1 + b	0
Be agnostic (abandon commitment)	M	M

FIGURE 5. Payoffs after Negative Evidence, with Commitment Benefits

So it will be beneficial to keep the commitment as long as  $p(X | \text{not-E}) > \frac{M}{1+b}$ , i.e., as long as  $p(X | \text{not-E})$  is sufficiently high.<sup>39</sup> This will hold under even weaker conditions if the initial benefits can only be realized by imposing costs, i.e., if committing to believe 'X' involves changing the  $t_2$  decision problem so that M is replaced with  $M - d$ :  $p(X | \text{not-E}) > \frac{M-d}{1+b}$ .<sup>40</sup> (Sometimes commitments have this form, when commitments are costly to break. An intention is a weak form of costly-to-break commitment, but there are also costlier-to-break commitments: resolutions, public declarations, and the like. Such a commitment might be worth it, given the payoffs at time  $t_1$ .)

<sup>39</sup>  $EU(\text{stick with commitment}) = p(X | \text{not-E})(1 + b) + p(\text{not-X} | \text{not-E})(0) = p(X | \text{not-E})(1 + b)$

$EU(\text{abandon commitment}) = p(X | \text{not-E})(M) + p(\text{not-X} | \text{not-E})(M) = M$

$EU(\text{stick}) > EU(\text{abandon}) \Leftrightarrow p(X | \text{not-E})(1 + b) > M \Leftrightarrow p(X | \text{not-E}) > \frac{M}{1+b}$ .

See also Fleisher (2018), whose *Inertia* condition requires that the individual's utility for continuing to endorse a proposition significantly increases after she initially endorses it.

<sup>40</sup> This will be additionally complicated by the evaluation of whether to change the decision problem. One difference between the cost of breaking a commitment ( $d$ ) and the mere absence of the benefit ( $b$ ) is how well one fares by being agnostic compared to how one would have fared by being agnostic if one hadn't made the commitment to believe in the first place: if the benefit merely goes away, one fares equally well; if breaking the commitment has costs, than one fares worse.

So we have the following result. If the following three conditions are met, then it is rational to commit to believe ‘X’, and to follow through on that commitment when one gets negative evidence, even though one wouldn’t have a high enough credence to believe ‘X’ after getting the negative evidence, if one hadn’t made the commitment:

- (i)  $p(X)$  is sufficiently high.
- (ii)  $p(X \mid \text{not-}E)$  is middling rather than low.
- (iii) There are benefits to be gained from committing to believe ‘X’, if  $X$  holds, that couldn’t be gained from merely believing ‘X’, and one must actually follow through on the commitment in order to get these benefits (or to avoid paying the costs of backing out).<sup>41</sup>

Thus, if there are benefits to committing now, *and* these benefits can only be captured if you follow through (or impose a cost to not following through), then it can be rational both to commit to believe ‘X’ and to follow through at a later time even if it would not have been rational to believe ‘X’ anew at the later time.

The reason for this result, put in intuitive terms, is this. When there is a high enough probability that negative evidence is misleading—would lead you to stop believing ‘X’ even though  $X$  holds—and there are benefits to committing and remaining committed to the belief if it is true, then the advantages of possibly shedding the belief if it’s false are outweighed by the advantages of possibly keeping the belief if it’s true. In other words, if there is still a sufficiently high probability that ‘X’ is true, then the advantages of sticking with a belief you’ve already ‘invested’ in, for the chance that it’s true, can outweigh switching course, for the chance that it’s false. The commitment doesn’t itself provide evidence for ‘X’; instead, given the benefits it confers on believing ‘X’ at  $t_2$ , it lowers the evidential bar needed to believe ‘X’ at  $t_2$

---

<sup>41</sup> Two quick numerical examples, to give the reader a sense of the magnitude of the values:  $M = 0.75$ ,  $p(X \& E) = 0.85$ ,  $p(X \& \text{not-}E) = p(\text{not-}X \& E) = p(\text{not-}X \& \text{not-}E) = 0.05$ ,  $p(X) = 0.9$ ,  $p(X \mid \text{not-}E) = 0.5$ , and either (i)  $b = 0.6$  OR (ii)  $b = 0.2$  and  $d = 0.2$ . Expected utilities: (i) At  $t_1$ ,  $EU(\text{commit}) = (0.9)(1.6) + (0.1)(0) = 1.44$ ,  $EU(\text{believe \& re-evaluate}) = (0.85)(1) + (0.05)(0.75) + (0.05)(0) + (0.05)(0.75) = 0.925$ ; at  $t_2$ ,  $EU(\text{stick with commitment}) = (0.5)(1.6) + (0.5)(0) = 0.8$ ,  $EU(\text{abandon commitment}) = 0.75$ . (ii) At  $t_1$ ,  $EU(\text{commit}) = (0.9)(1.2) + (0.1)(0) = 1.08$ ,  $EU(\text{believe \& re-evaluate}) = (0.85)(1) + (0.05)(0.75) + (0.05)(0) + (0.05)(0.75) = 0.925$ ; at  $t_2$ ,  $EU(\text{stick with commitment}) = (0.5)(1.2) + (0.5)(0) = 0.6$ ,  $EU(\text{abandon commitment}) = 0.75 - 0.2 = 0.55$ . Note that in the first example, given that  $p(X \mid \text{not-}E)$  is fairly low, the benefits of commitment must be very high to meet the conditions  $(0.5b > M - 0.5)$ : high enough that for many initial values of  $p(X)$ , both committing to believe ‘X’ and committing to believe ‘not-X’ will be preferable to agnosticism.

(without also lowering the evidential bar needed to believe ‘not- $X$ ’). It provides a reason to continue to believe ‘ $X$ ’ even though one would not believe ‘ $X$ ’ without the commitment.

The result here goes beyond the obvious point that if there are benefits to commitment, then it can be good to maintain a belief even if it wouldn’t have been good to believe it in the absence of these benefits. What the model contributes is showing which things, exactly, we ought to be weighing against each other, and how the magnitude of the new probability matters. And, more importantly, it shows that benefits to commitment can make it rational to maintain a belief in the face of doubt, even if those benefits only obtain if the belief is true.

One final note. Although the view here recommends maintaining a commitment in the face of negative evidence, it does not recommend doing so indefinitely, or in the face of any negative evidence whatsoever. (Commitments are not usually vows.) If evidence against one’s view mounts (i.e. if credence continues to decrease) or if credence is middling for a long enough time to eclipse the benefits of commitment, then it may be rational to become agnostic.

## 5. The Benefits of Committing to the Truth

This result suggests a distinction between what we might call ‘ordinary belief’ and ‘conviction’—the latter being a belief that we not only hold but are committed to continue holding (perhaps: have the intention to continue holding) even in the face of a certain amount of counterevidence. And it is rational to have convictions, rather than mere ordinary beliefs, if we are in a certain epistemic situation— $p(X)$  is high and  $p(X | \text{not-}E)$  is middling rather than low—and if there are benefits to *committing* to believe ‘ $X$ ’, if  $X$  turns out to be true, beyond just the benefits of believing ‘ $X$ ’.

When, if at all, does this latter condition hold? That is, what sort of benefits might be gained from committing to believe ‘ $X$ ’, if  $X$  is true, that couldn’t be gained from merely believing ‘ $X$ ’? I want to set aside cases in which the benefits are unrelated to the truth of the proposition (e.g. someone threatens to torture you if you don’t commit to believe ‘ $X$ ’), and focus on benefits I will call *quasi-epistemic*, that is, benefits with two features. First, they only obtain if the claim believed is true. And second, they are considerations which could either directly move us to believe something or could plausibly be the result of some psychological mechanism that could move us. This section will discuss possible benefits along these lines.

The first possibility is that, on particular matters of epistemic importance, one prefers to continuously believe the truth over any other pattern of belief. That is, one thinks it better to be unwavering in one's belief in the truth rather than to sometimes believe the truth and sometime remain agnostic, independent of where the evidence points at a particular time. This will hold only in domains in which we have a particular concern with this pattern of correctness. One such domain appears to be the domain of interpersonal relationships. We might think it better, for example, to always believe that our friends are trustworthy, if they in fact are, than to sometimes be agnostic. Insofar as religion contains an interpersonal element (one's relationship to God or to one's religious community), religious beliefs are like this as well. A case could be made that some moral or political beliefs involve a relationship to a community as well, and thus share this feature.

Another benefit that might be gained from committing to believe a proposition is what we might call the benefit of *confident evaluation*. When we are committed to a belief, we shape our reasoning according to its truth. We cut down on the number of options.<sup>42</sup> We ignore possibilities that do not make sense if the belief is true, and rule out sources that reason from not-*X*, thus cutting down on our cognitive costs. We collaborate with others who also believe '*X*', to gain evidence. We pay more attention to propositions whose truth follows from that belief, i.e., we try to flesh out a theory which has as its basis the truth of *X*. For example, if we take consequentialism to be true, then we devote resources to figuring out how to do the most good possible (as in the effective altruism movement), and we don't spend time debating what follows from the categorical imperative or how to understand the classical virtues. It is useful to commit to a claim that helps shape and sift our reasoning—to proceed in our reasoning on the supposition that this claim is true and ignore the possibility that it is false—and it is expensive to drop this commitment, since one has then to start much of one's reasoning from scratch.<sup>43</sup>

---

<sup>42</sup> See Holton (2004, 2014), although he primarily focuses on the costs of reconsidering the evidence, not on the costs of changing course or changing views once one has reconsidered. (He would say: once we're reconsidering, start anew. But I would say: even when reconsidering, sometimes stay the course.) Here we possibly depart from Levi's model, since his is more holistic in nature, e.g., evaluating all of the atomic propositions one cares about. A similar point to the above may be modelable on his picture, but I won't consider details here.

<sup>43</sup> Others have noticed that individuals, each committed to their own view, can effectively divide cognitive labor, thus serving our collective goal of seeking truth. See Elgin (2010), Fleisher (2018), Loughheed (2018a), and Kraay (2019) on disagreement in particular, as well as classic papers by Kitcher (1990), Strevens (2003), and Muldoon (2013). But here I note the benefits to the individual, if *X* holds.

How does such a commitment co-exist with credence, and in particular, credence that changes in rational ways? While it is useful to ‘anchor’ aspects of one’s cognitive life with a claim that one takes to be true, and to remain committed to that claim for purposes of reasoning and searching for evidence, such commitment cannot be absolute. Thus one also needs to keep track of one’s credence in the claim, so that one can drop the commitment when one’s credence in  $X$  drops low enough that the benefits of commitment (if  $X$  holds) become unlikely enough to obtain, and thus do not outweigh the detriments of believing what is likely to be false. The committed consequentialist must be committed, but she should also be tracking how she assesses the evidence for her theory, so that she can drop her commitment if her evidence dips below a (fairly low) threshold. The picture here can thus point to a way in which both credence (gradational assessment of truth-value) and belief (on-off assessment of truth-value) are necessary parts of cognitive life: the former tries to track strength of evidence ‘objectively’ or ‘from nowhere’, and the latter responds to the evidence but can also be sensitive to the shape of one’s cognitive life over time.

Notice that the benefits of confident evaluation are more apt to obtain when ‘ $X$ ’ is a *high-level proposition*, i.e., when ‘ $X$ ’ is deeply embedded in the web of things we believe, when lots of things follow from it, or when what we accept informs which evidential sources we consult.<sup>44</sup> The most extreme case here is when the proposition in question forms the core of a paradigm in the sense of Kuhn, or the core of a research program in the sense of Lakatos.<sup>45</sup> And, presumably, high-level propositions are exactly the kind of propositions at issue in many disagreements about morality, politics, and religion.<sup>46</sup>

---

<sup>44</sup> See also Bishop (2007), who argues that having faith involves taking a high-level claim, for which the evidence is inconclusive, to be true for practical purposes. The picture here says: faith involves believing a claim for which you already have evidence, and sticking with it in the face of counterevidence; and such faith is apt to be rational when the claim is high-level.

<sup>45</sup> Kuhn (1970), Lakatos (1970). The question of assessing disagreement between paradigms or research programs—and in particular how to assign credences to core propositions—is beyond the scope of this paper. One difficulty is that paradigms are often thought to be ‘incommensurable’, so it is not clear in what sense two committed proponents of different paradigms are ‘disagreeing.’ Nonetheless, I take it for granted that in the cases in this paper proponents of different views can assess the other’s view and also assign credence to the other’s hypothesis. (For example, the theist understands what the atheist is saying, and assigns an initially low credence to atheism.)

<sup>46</sup> Indeed, notice that TENACITY becomes less compelling when we are talking about ‘lower-level’ propositions within a given framework. It seems more compelling that a rabbi ought to stick to his belief in Judaism in the face

A final benefit that might be gained from committing to believe a proposition we might call the benefit of *confident action*. Assume a picture according to which one acts on what one believes. If one is engaged in a ‘long-term project’ on the basis of a claim—a project that consists in a series of actions over time each of whose rationality depends on believing the claim in question—then under the assumption that this claim is true, it is better to believe the claim at every time than to have any other pattern of belief, so that one can successfully perform the series of actions. Translated to the model here: there is a benefit that accrues to ‘commit to believe ‘X’ if *X* holds because, assuming we act on what we believe, it is sometimes advantageous to be committed to act on a true claim even in the face of doubts, since certain projects require action over an extended time.

Many religious projects are like this. For example, being a nun requires being a continuous part of a religious community, and even ordinary religious devotion typically requires rituals or actions that are repeated over time. If the claim upon which the religious devotion is based is true, then it is better to show up every week and perform the rituals or actions than it is to disappear for stretches when one’s credence is lower—even rationally lower. Moral projects are like this, too. Helping to alleviate poverty in a particular community, solving some social problem, and working towards a cure for some illness each require work over time, and this work is usually based on an uncertain claim—a claim which is the object of considerable disagreement—about whether this particular devotion of time and resources will be effective. Again, if it is in fact effective, then putting time and resources in steadily is better than varying them based on one’s current credence level. (Even if such action isn’t *based* on belief, it seems like belief helps in that it allows one to keep one’s will steady in the face of uncertainty.)

An alternative picture of action holds that we act on our *credences*, and perhaps in addition—contra the Levi view—that we count as believing a claim insofar as we are prepared to act on it given the stakes.<sup>47</sup> I will not discuss this picture in detail here, but we will get a similar result: we ought to commit to act on a claim if the above conditions hold, because there are benefits to staying committed to a long-term project, namely that that remaining committed is the only possibility for completing it.<sup>48</sup>

---

of disagreement about its truth, than that two rabbis ought to each stick to their beliefs in particular interpretations of what the Jewish faith says when disagreeing with each other.

<sup>47</sup> Views in this general spirit include those of Harsanyi (1985), Fantl and McGrath (2002), Weatherson (2005), and Pace (2011).

<sup>48</sup> See Buchak (2017).

If a conviction is a belief we are committed to, even in the face of counterevidence that would make us not form the belief anew, then when the above credence conditions are in place—when one’s current credence in a claim is high and one’s conditional credence in that claim given potential counterevidence is middling rather than low—and the belief concerns a proposition for which there are commitment benefits, it can be rational to adopt a conviction and to maintain that conviction when counterevidence arises. It can be rational to maintain a conviction even if one’s credences would not make it rational to adopt that belief anew. (This is surprising, and, again, provides an interesting way in which credence and something like belief come apart: credence tracks the evidence at a time, and conviction is a more stable phenomenon.<sup>49</sup>)

To be clear, my conclusion isn’t that one should avoid new evidence because, since ‘X’ is true, evidence against *X* must be misleading. (I am not endorsing Kripkean dogmatism.) Instead, *one should update one’s credences on the counterevidence, but one should maintain one’s belief despite the counterevidence*. Given that one’s credence is now middling—i.e. there is still a good chance that *X* is true—the possibility that the evidence is misleading, coupled with the benefits of (some chance of) continuing to believe the truth means that the benefits of agnosticism are outweighed by the benefits of continuing to believe ‘X’.

The picture here has several interesting upshots. First, if there are commitment benefits regarding both *X* and not-*X*, then the following situation might arise. I might start with a conviction that *X* and you might start with a conviction that not-*X*, both rationally based on appropriate credences given our evidence. We might then both encounter new evidence (or perhaps share evidence with each other), such that we both end up with a credence in *X* of 0.5. However, it might still be the case that each of us should remain committed to our initial belief, despite now sharing the same credence. What is rational to believe depends on our initial convictions.

Second, the picture here provides an argument for a limited kind of conservatism in belief, without conservatism in credence. When it comes to important or fundamental claims, one ought to stick with one’s current beliefs even as evidence comes in that makes these beliefs no longer rational to adopt for someone who doesn’t already hold them. This picture also explains why a belief-conservationist might be

---

<sup>49</sup> See also Holton (2014) on the stability of belief, although he argues against the idea that we humans have credences at all.

keen to include credence in their ontology: a conservatism about belief that didn't tell us precisely when to give up the belief would be more difficult to defend.

## 6. Vindicating FREQUENCY, HUMILITY, and TENACITY

In the previous section, the following fact about evidence, credence, and belief emerged:

FACT: When new evidence comes in, it can be epistemically rational to update one's credence, while simultaneously remaining committed to a belief that one would not adopt based on one's new credence if one didn't already hold that belief. The circumstances in which this is apt to be rational are that the initial balance of the evidence tells in favor of the relevant claim, counterevidence does not tell strongly against the claim, and there are benefits to remaining committed to the belief if it is true.

Let us return to the topic of disagreement. We are now in a position to vindicate FREQUENCY, HUMILITY (EVIDENCE), and TENACITY (BELIEF), precisely because in certain circumstances—those described in FACT—SYNCHRONICITY is false. (Recall, SYNCHRONICITY is the claim that my belief after I encounter peer disagreement should depend only on the balance of my evidence after I encounter peer disagreement.) We will do this by assuming that FREQUENCY and HUMILITY (EVIDENCE) are true, and show that TENACITY (BELIEF) comes out true as well.

Recall the conditions:

FREQUENCY: There is widespread disagreement among peers about the DISPUTED CLAIMS.

HUMILITY (EVIDENCE): When an epistemic peer forms a very different opinion about a claim than I do, the balance of my evidence for that claim changes significantly.

TENACITY (BELIEF): I ought to maintain my belief about the DISPUTED CLAIMS.

To flesh out HUMILITY (EVIDENCE), we will make a few assumptions. The first is that credence aims to reflect the balance of the evidence at a time. The second is that one ought to respond to disagreement by conditionalizing on the fact that the individual with whom one disagrees has a particular credence in  $X$  or particular belief in the matter of  $X$ . As Adam Elga puts it:

“Upon finding out that an advisor disagrees, your probability that you are right should equal your prior conditional probability that you would be right. Prior to what? Prior to your thinking through the disputed issues, and finding out what the advisor thinks of it. Conditional on what? On whatever you have learned about the circumstances of the disagreement.”(Elga 2007: 490)<sup>50</sup>

Thus, encountering a disagreeing peer is a special case of receiving counterevidence more generally, and peer disagreement in which participants have exactly the same evidence is continuous with the case in which they simply believe each other to be in roughly equally good evidential positions.

The final assumption is that one’s conditional credence in ‘*X*’, given that a peer forms a very different view, typically ought to be substantially different from one’s credence in ‘*X*’ prior to learning of the disagreement. (Thus, we will assume a version of conciliationism for credence.) An example of a view that meets these conditions is the view that I ought to treat my opinion and that of my peer symmetrically—in the sense of thinking we are equally likely to get it right—and respond to the disagreement by forming the same credence I would form if our opinions were reversed.<sup>51</sup>

If we adopt these assumptions, then if I meet a peer with a very different credence from mine, I (typically) ought to drastically alter my credence. Thus, when I form a high credence in *X* and meet a peer with a low credence in *X*, I will typically end up with a middling credence—not high, because the peer’s credence is low, and not low, because my initial credence is high. Therefore, if we take the existence of the disagreeing peer to be the relevant counterevidence and if we assume I have an antecedently high

---

<sup>50</sup> See also Pittard (2020: 110).

<sup>51</sup> This doesn’t necessarily mean that I ought to ‘split the difference’ between our credences in the sense of averaging. As Christensen (2009) points out, splitting the difference both runs into technical difficulties (here, he cites Shogenji (2007) and Jehle and Fitelson (2009)) and belies the motivation for conciliationism: for example, if two individuals independently form a high credence in *X* then this is evidence that they ought to be even more confident in *X*. (See also Nissan-Rozen and Spectre (forthcoming).) We can think of ‘symmetry’ conciliationism as the view that two or more disagreeing peers ought to each adopt the credences that would be licensed by pooling their opinions. (For a sampling of literature on credence pooling, see Lehrer and Wagner (1983) and Russell et al (2015).) It is possible that in some cases ‘pooling’ would force both one and one’s peer to adopt one’s own initial credence, but only by happy accident. See also Pittard (2020: 101-107) on the role that prior probabilities play when peers are treated as ‘equally reliable instruments’.

credence in 'X', then conditions (i) and (ii)— $p(X)$  is high and  $p(X \mid \text{not-E})$  is middling rather than low—are met in the situation of peer disagreement.

I've argued above that claims about morality, religion, and politics are particularly apt to meet condition (iii), the condition that there are benefits to committing to these claims as long as they are true. To recall, this is for one of several reasons: there are benefits to continuously believing the truth over and above the benefits of believing the truth at particular times; the claim is a 'high-level' claim which figures deeply into one's web of reasoning; or the claim forms the basis of a long-term project.

Given these facts, if I meet a peer who disagrees with me about an important claim in one of these domains, it is rational to lower my credence in the claim but remain committed to believing the claim. Therefore, TENACITY (BELIEF) is satisfied. It is satisfied because (contra SYNCHRONICITY) my beliefs after I encounter peer disagreement depend partially on my *prior* evidence--on the commitments that my assessment of this evidence supported--not just on my *new* total evidence, i.e., not on my assessment of the evidence which includes the existence of the disagreement.

To illustrate, assume an individual who has a high credence that consequentialism is true meets an epistemic peer with a low credence that consequentialism is true; assume further that the first individual's credence is high enough for her to rationally believe in consequentialism, and the second individual's credence is low enough for him to rationally disbelieve in consequentialism. The consequentialist ought to take the existence of the disagreeing peer to be evidence against her own opinion: thus, she ought to lower her credence in consequentialism. But she knows that there is a decent chance that the evidence is misleading, and she has more to lose than just a true belief if she gives up her belief in consequentialism and it turns out to be true. And while she doesn't know that *this evidence* is misleading—which is why she genuinely ought to lower her credence—she knows that the only way to reap the benefits of believing in consequentialism over the long run is to sometimes stick it out during periods of doubt.

Thus, we can give voice to the idea that I shouldn't privilege my opinion over yours: as far as my new credence is concerned, it needn't make a difference whose opinion was whose. And we can simultaneously give voice to the idea that there is something special about *my own* beliefs about DISPUTED CLAIMS: I can maintain them even in the face of my new credence; indeed, there can be situations in which you and I end up with the same credences but both retain our initial beliefs. We can be conciliationist about our *credences*, and steadfast about (some of) our *beliefs*.

(Notice that any theory that can justify a minimal conservatism—that allows one to retain one’s beliefs for a period of time even as one’s evidence changes—may have the resources to respond to the problem of disagreement by denying SYNCHRONICITY.<sup>52</sup> So if one is dissatisfied by the story here but wants to retain the three claims outlined in section 2, then this is an avenue to explore further.)

The recommendation here will lead to some tension: my current beliefs will be irrational according to my new total evidence, setting aside the fact that I am already committed to them. I will be in the uncomfortable position of believing something that I no longer take to be rational from a ‘view from nowhere’—the uncomfortable position of believing something that I would not be rational to believe if I approached the question anew. Perhaps I will even wish that I hadn’t made the commitment in the first place. But this tension is in fact felt when we respond to disagreement by maintaining our own beliefs, and so it is an advantage of the present view that it can explain this tension.

According to views which maintain TENACITY by circumscribing FREQUENCY, I simply do not consider those who disagree with me on important matters to be peers. According to many views which maintain TENACITY by denying HUMILITY, I can locate some feature which makes my situation superior. On both of these kinds of views, it seems that I shouldn’t feel any tension—or not more tension, at any rate, than an expert would feel if a novice disagreed with her. But the truth is that when we encounter a thoughtful, knowledgeable person who disagrees with us about something important, we usually do (and, I think, ought to) feel some tension in continuing on as before. And that is precisely what the view here predicts. According to the view here, we ought to continue on as before, but not because of or by following our evidence at the current time; rather, by honoring past commitments, by staying the course. We ought to continue on while acknowledging that the existence of the peer makes our own convictions or actions unjustified from a ‘universal’ perspective.

One might ask whether accepting the tension between my beliefs and the total evidence is itself an unacceptable way of circumscribing TENACITY. After all, isn’t it sometimes important to maintain our beliefs not just about claims themselves, but about what the evidence supports?<sup>53</sup> It might be important to

---

<sup>52</sup> A view in this neighborhood is Sarah Paul’s view (in Paul (2015)) that one ought to avoid reassessing the evidence when (and because) reassessing the evidence is a kind of temptation. Paul holds that this is true when (and because) one knows one is being irrational: one knows that non-truth-conducive factors (wishful thinking and the like) color one’s current assessment.

<sup>53</sup> I thank an anonymous referee for this objection, including the specific examples.

me not just to believe that climate change is human-caused, or that God exists, but also to believe that *the evidence supports* these claims. However, while it might make one rest easier in one's beliefs about climate change and God's existence if one could believe those claims about the evidence, they nonetheless aren't typically claims about which it is important to have *convictions*: one's pattern of belief about them does not seem important, and they do not generally form the basis of other reasoning or of action. (Insofar as it is, or they do, then one may be able to have convictions about them—but then the tension will reappear elsewhere, since one's beliefs about the evidence won't be supported by one's current evidence about the evidence.) Alternatively, it might be important to maintain my beliefs about the evidence because consistency is important to me: I want to be the type of person whose beliefs about things only when I have an active belief that the evidence (considered objectively, and only at a time) supports them. In response, I can only say this: if this is important, and if one is loath to give up HUMILITY, then the only way to respond to disagreement is by becoming agnostic about many things. This may be acceptable to some, but what I hope I've shown is that there are good reasons to hold onto one's beliefs, even at the expense of some tension—and that doing so is more rationally laudible than one might have thought.

## 7. Cases

Let us explore what verdicts the view in this paper delivers when applied to some of the other kinds of cases in the disagreement literature: mundane cases, disagreement with many people, disagreement with one's possible self, and disagreement known prior to forming a view of one's own.

First, consider mundane cases such as David Christensen's mental math case.<sup>54</sup> In this case, you are dividing a bill evenly among friends, and you and a peer each arrive at a slightly different answer about how much each person owes. Most philosophers writing on the topic seem to agree that you should immediately become less confident of your own answer (cease believing it, or lower your credence, depending on the attitude type under discussion). And, happily, that is what the view here enjoins as well. There is no reason, before hearing the other person's calculation, to adopt a conviction that your calculation is correct: you don't care that much about maintaining a correct belief on this topic over time, you don't have a lot of other beliefs that depend on presupposing this calculation in your reasoning, you are not already involved in a long-term course of action on its basis, and so forth. Thus, you should lower your credence in your own answer, and also become agnostic.

---

<sup>54</sup> Christensen (2007).

A lot depends on the specific details of the case, so I cannot issue a blanket pronouncement about mundane cases. But I suspect that many mundane cases have this character: they concern claims whose truth one is only interested in temporarily, that do not form the core of a web of reasoning, and that do not undergird long-term projects. This contrast between mundane and significant cases recalls the defense that some have given for circumscribing FREQUENCY so that we can retain both HUMILITY and TENACITY: the defense that some beliefs are so embedded in your web of beliefs that you cannot consider others peers with respect to them.<sup>55</sup> The explanation in the previous section recommends tenacity when claims are so embedded, but for a very different reason. You still must consider others peers and therefore modify your assessment of the balance of the evidence (your credence), but your *beliefs* about claims that are embedded shouldn't be tied to the new balance of the evidence.

Second, consider cases in which one forms an opinion and then encounters *many* people who disagree, and the corresponding question of whether 'numbers matter.' Recall that on the view here, beliefs of the appropriate type should be tenacious with respect to lowered credence, but should not be tenacious with respect to drastically lowered credence: when one's credence is drastically lowered, or perhaps when it stays moderately low for a long period of time, one ought to rethink one's previous belief commitments. Thus, if one encounters many people who disagree with one (and none who agree), and their opinions are genuinely independent from each other (but note that this is more likely to happen in, say, perceptual cases than political cases), then one ought to drastically lower one's credence and rethink one's commitments. Thus, convictions provide a kind of inertia, but one that can be overcome.

Third, consider disagreement with one's possible self. For example, let's say you have been trained in a graduate program where everyone accepts consequentialism and you come to be a consequentialist, but you reasonably believe that if you had gone to a different, equally good graduate program, you would have become a deontologist.<sup>56</sup> Or, you might take a particular religious or political view to be broadly correct, but hold that had you grown up in a different location, you would likely be just as confident about a different view. The picture here allows us to hold that this counterfactual disagreement should make you lower your credence in the relevant claim—it should change how you see the balance of the evidence, that it no longer strongly points to your own view—but that nonetheless you are entitled to continue on

---

<sup>55</sup> Elga (2007). See related points in Pettit (2005) and Pittard (2014) and challenges by McGrath (2008), Loughheed (2018a), Hallsson (2019), Fritz and McPherson (2019), and Fritz (2018).

<sup>56</sup> See Kelly (2010), who cites Cohen (2000), for a similar example.

with your convictions. In other words, had your circumstances been different, you would have (rationally) adopted a very different set of commitments, and these would be currently guiding your beliefs—but in your circumstances you are entitled to fulfill the commitments you actually have, even while your credences reflect that neither set of circumstances is more truth-conducive.

Finally, consider cases in which one already knows what other people think. In many—maybe most—cases of religious, moral, and political beliefs, we know ahead of time that whichever conclusion we come to, there are peers who will disagree with us. Although full discussion of these cases would take us too far afield, there's a lesson here for them as well, if we assume that our own assessment of the evidence should be *conceptually* prior to our assessment of the evidence when we include other's opinions. In particular: if it is rational to commit to a claim on the basis of this smaller subset of evidence (the evidence excluding others' opinions), then the total evidence might make it rational to maintain that commitment if the relevant conditions apply.

## 8. Principles

Finally, we turn to implications of the view here for particular principles and debates in epistemology.

We turn first to what is called the Uniqueness Thesis: the thesis that a given body of evidence licenses a single doxastic attitude.<sup>57</sup> The Uniqueness Thesis closes off one way to reject HUMILITY, namely rejecting it because facts about the believer herself are relevant to what doxastic attitude she should form on the basis of her evidence (so she need not treat herself and her peer symmetrically). The view in this paper is compatible, for all we've said, both with accepting and rejecting Uniqueness about credences.<sup>58</sup> But interestingly enough, on the view here, one can *accept* Uniqueness about credences but *reject* Uniqueness about beliefs: one can hold that a given body of evidence (at a time) doesn't entail an answer to what one's beliefs should be—even holding fixed the utility of a true belief—because what one's beliefs should be depends not just on one's current evidence, but on one's history or past commitments.

The discussion here also bears directly on the principle that my epistemic state should not depend on the order in which I acquire evidence. This is articulated by Richard Feldman:

---

<sup>57</sup> Feldman (2007), Christensen (2009), Cohen (2013).

<sup>58</sup> For discussion of the relationship between Uniqueness and Conciliationism, see Kelly (2010) and Ballantyne and Coffman (2010).

“It is worth adding that the order in which one gets one’s evidence on the topic makes no difference. . . . The principles of rational belief do not include a law of inertia.”(Feldman 2007: 10)

And by Thomas Kelly:

“It seems implausible (to say the least) that historical facts about the order in which evidence is acquired might make such a dramatic difference to what one is justified in believing. Indeed, many take it to be a criterion of adequacy on any account of rational or justified belief that the order in which pieces of evidence are acquired makes *no difference at all* to what is reasonable to believe. This is the frequently endorsed requirement that evidence be commutative” (Kelly 2008: 6-7)

Again, if we accept the idea defended here—that the rationality of belief can depend both on one’s current evidence and on what it was rational to commit to believe in the past—then this claim is false. In certain cases, the order in which I receive evidence will (rationally!) make a difference to which beliefs I should hold. An individual who forms a belief that ‘X’ and her peer who forms a belief that ‘not-X’ might walk away from a disagreement with their credences aligned but their original beliefs intact, provided each belief played the right role in each individual’s epistemic life. What it is rational to believe can be path-dependent.

Would accepting the view here therefore force us to rethink much of what is sacred in epistemology? No. First, the epistemology of credence remains unchanged: the balance of the evidence, and my response to it, will depend in the ordinary way on whatever epistemologists conclude it depends on.<sup>59</sup> Second, the question of how to fit on/off belief into the framework of credence was already a vexed question, and it is simply an interesting result that they have different roles and are governed by different norms. Finally, the phenomenon of retaining one’s beliefs by ‘ignoring’ new evidence only occurs in a limited range of cases: when one has a particular credal history, the evidence has a particular character, and the belief concerns a claim of the relevant type.

---

<sup>59</sup> Note, incidentally, that a more general form of conditionalization (Jeffrey conditioning) violates the commutativity of evidence (see Kelly (2008: 7, footnote 8)).

I've argued for a 'split-attitude' picture, in which one is 'conciliationist' with respect to one kind of epistemic attitude, and 'steadfast' with respect to another kind. A different split-attitude picture is Allan Hazlett's, according to which you should steadfastly retain your first-order doxastic attitude toward the disagreed-upon proposition, but adopt a conciliatory second-order attitude about the epistemic status of your first-order doxastic attitude.<sup>60</sup> On the view here, the split is not between what one should do with respect to one's first-order attitude and second-order attitude—with respect to different attitude 'levels.' Rather, the split is between what one should do with respect to different attitude *types*. Both your credence in 'X' and belief that 'X' are first-order attitudes towards 'X', and one should be conciliatory with respect to credences and steadfast with respect to (some) beliefs.

In conjunction with this point, it bears mentioning the phenomenon of *epistemic akrasia* and its debated rationality. The akratic agent believes 'X' but also believes that the evidence does not support 'X', or (in a weaker form) believes 'X' but suspends judgment about whether the evidence supports 'X'.<sup>61</sup> One might naturally ask, is the individual who follows the course I have laid out here being akratic, and if so, is this a problem? It depends on how we understand 'the evidence supports', in two ways. First, is the relation taken by the subject to be universal and objective—to mean something like, 'anyone with this evidence ought to believe *X*'—or to be relativized to her situation—to mean 'someone with my commitments and circumstances and this evidence ought to believe *X*'? Second, are we taking evidential support to be synchronic—'the evidence supports *X*'—or diachronic—'the newly acquired piece of evidence isn't sufficient to remove support from *X*'?

The subject's commitment lowers the bar for how much evidential support (how high a credence) is needed for her to believe at a later time, and it lowers this bar for reasons concerning the subject's initial evidence. At the later time, the evidence considered universally or objectively (i.e. without her own commitment in view) does not support *X*. But the evidence relativized to her cognitive commitments does support *X* (because of the lowered bar). Furthermore, if we consider the evidence universally or objectively, the evidence at the prior time supported *X* and the newly acquired piece of evidence isn't sufficient to remove support from *X*. Finally, the agent can be aware of all of these facts. So, we might say that the subject (at least, if she is aware of all the facts) has a limited kind of akrasia: she believes 'X'

---

<sup>60</sup> Hazlett (2012). His paper focuses on belief, rather than credence, but I don't think this is crucial to his view. See also Turnbull and Sampson (2020).

<sup>61</sup> See, for example, Owens (2002), Coates (2012), Hazlett (2012), Horowitz (2014), Lasonen-Aarnio (2014), and Lasonen-Aarnio (2018).

while being aware that not everyone with her current evidence ought to believe ‘X’, or while being aware that the evidence does not objectively and synchronically support *X*. If this is right, then it seems like we’ve identified a new case in which a subject might exhibit a kind of akrasia, and a reason for it.

Some of the debate over akrasia focuses on whether being in a state of akrasia is possible. If the subject here does count as akratic, it is easier to see how akrasia is possible in this case. While many standard cases focus on whether a subject can *form* a belief when she does not take her evidence to warrant it, here we have a case in which a subject *maintains* a belief when she no longer takes her evidence to warrant it (more accurately: she maintains a belief when she holds that her evidence would not support forming a belief if she were to approach the question anew). The supposition that we have psychological mechanisms that keep us believing despite evidence that calls that belief into question—and transparently so—is much weaker than the supposition that we have psychological mechanisms that allow us to actively form a belief despite such evidence.

Whether epistemic akrasia is possible or rational, therefore, may depend on whether evidential support is synchronic or diachronic. If it is diachronic, then we do not have here a case of akrasia, because the subject is correctly relating her current beliefs to her old beliefs, given the intervening evidence. But if it is synchronic, then akrasia is both possible and rational because the mechanisms of belief formation and maintenance, and the norms of rationality, are not themselves synchronic.

A final way in which the argument here has implications for debates in epistemology, as already discussed, is that it can explain the pull of a limited kind of conservatism with respect to significant beliefs. People in fact do exhibit this kind of conservatism. If before we thought this was irrational or exhibited a kind of prejudice, we now have a story according to which it can be vindicated.

## 9. Conclusion

Many peers disagree with you about important claims, and peer disagreement can significantly shift the balance of your evidence. But your beliefs at a given time needn’t always depend on the balance of your evidence at that time. On the contrary, in some cases there are important reasons not just to believe a claim, but to have a *conviction*—to commit to believing the claim even in the face of a certain amount of counterevidence—and then to be guided by this conviction when the balance of the evidence changes. Convictions can be beneficial when a pattern of continually believing the truth with respect to a claim is better than any other pattern, or the claim shapes and sifts your reasoning, or the claim forms the basis of an important long-term project. When it comes to claims with one of these features, you can and should

continue believing what you believe, even upon discovering that others disagree with you. And so even though disagreement significantly shifts the balance of your evidence, you can remain steadfast in many of your beliefs about important matters.

## BIBLIOGRAPHY

- Ballantyne, Nathan & E. J. Coffman (2012), 'Conciliationism and Uniqueness', *Australasian Journal of Philosophy* 90(4): 657-670.
- Barnett, Zach (2019), 'Philosophy Without Belief', *Mind* 128(509): 109-138.
- Bergmann, M. (2017), 'Religious Disagreement and Epistemic Intuitions', *Royal Institute of Philosophy Supplement* 81: 19-43.
- Bishop, John (2007), *Believing by Faith* (Oxford University Press).
- Bogardus, Tomas (2013), 'Disagreeing with the (religious) skeptic', *International Journal of Philosophy of Religion* 74: 5–17.
- Buchak, Lara (2010), 'Instrumental Rationality, Epistemic Rationality, and Evidence-Gathering', *Philosophical Perspectives* 24: 85-120 (2010).
- Buchak, Lara (2012), 'Can it be rational to have faith?', in Jake Chandler and Victoria S. Harrison (eds.), *Probability in the Philosophy of Religion*, 225-247. Reprinted in Louis P. Pojman and Michael Rea (eds.), *Philosophy of Religion: An Anthology, 7th edition* (Wadsworth, 2014).
- Buchak, Lara (2013), *Risk and Rationality* (Oxford University Press).
- Buchak, Lara (2014), 'Rational Faith and Justified Belief', in Tim O'Connor and Laura Goins (eds.), *Religious Faith and Intellectual Virtue* (Oxford, 2014), 49-73.
- Buchak, Lara (2017), 'Faith and Steadfastness in the Face of Counter-evidence', *International Journal for Philosophy of Religion* 81(1): 113-133.
- Carter, J. Adam (2018), 'On Behalf of Controversial View Agnosticism', *European Journal of Philosophy* (4):1358-1370.
- Coates, Allen (2012), 'Rational Epistemic Akrasia', *American Philosophical Quarterly* 49(2): 113-24.
- Choo, Frederick (forthcoming), 'The Epistemic Significance of Religious Disagreements: Cases of Unconfirmed Superiority Disagreements', *Topoi* 1-9.
- Christensen, David (2007), 'Epistemology of Disagreement: The Good News', *Philosophical Review* 116(2): 187-217.
- Christensen, David (2009), 'Disagreement as Evidence: The Epistemology of Controversy', *Philosophy Compass* 4/5: 756-767.
- Cohen, G.A. (2000), *If You're an Egalitarian, How Come You're So Rich?* (Cambridge, MA: Harvard University Press).
- Cohen, Stewart (2013), 'A Defense of the (Almost) Equal Weight View', in David Christensen and Jennifer Lackey (eds.), *The Epistemology of Disagreement: New Essays* (OUP): 98-117.

- Conee, Earl (2010), 'Rational Disagreement Defended', in Richard Feldman and Ted Warfield (eds.), *Disagreement* (OUP): 69-90.
- Dormandy, Katherine (2018), 'Resolving Religious Disagreement: Evidence and Bias', *Faith and Philosophy* 35(1): 56–83.
- Elga, Adam (2007), 'Reflection and Disagreement', *Noûs* 41(3): 478-502.
- Elgin, Catherine Z. (2010), 'Persistent Disagreement', in Richard Feldman and Ted Warfield (eds.), *Disagreement* (OUP): 53-68.
- Elgin, Catherine Z. (2018), 'Reasonable Disagreement', in Casey Rebecca Johnson (ed.), *Voicing Dissent: The Ethics and Epistemology of Making Disagreement Public* (Routledge, New York): 10-21.
- Fantl, Jeremy and Matthew McGrath (2002), 'Evidence, Pragmatics, and Justification', *Philosophical Review* 111(1): 67-94.
- Feldman, Richard (2007), 'Reasonable Religious Disagreements', in Louise Antony (ed.), *Philosophers Without Gods: Meditations on Atheism and the Secular Life* (Oxford University Press): 194-214.
- Fleisher, Will (2018), 'Rational Endorsement', *Philosophical Studies* 175: 2649-75.
- Fritz, James (2018), 'Conciliationism and Moral Spinelessness', *Episteme* 15(1): 101-118.
- Fritz, James and Tristram McPherson (2019), 'Moral Steadfastness and Meta-Ethics', *American Philosophical Quarterly* 56(1): 43-55.
- Fumerton, Richard (2010), 'You Can't Trust a Philosopher', in Richard Feldman and Ted Warfield (eds.), *Disagreement* (OUP): 91-110.
- Goldberg, Sanford (2009), 'Reliabilism in Philosophy', *Philosophical Studies* 142(1): 105-117.
- Goldberg (2013a), 'Disagreement, Defeat, and Assertion', in David Christensen and Jennifer Lackey (eds.), *The Epistemology of Disagreement: New Essays* (OUP): 167-89.
- Goldberg (2013b), 'Defending Philosophy in the Face of Systematic Disagreement', in Diego E. Machuca (ed.), *Disagreement and Skepticism* (Routledge): 277-294.
- Goldberg, Sanford (2014), 'Does Externalist Epistemology Rationalize Religious Commitment?', in Laura Frances Callahan and Timothy O'Connor (eds.), *Religious Faith and Intellectual Virtue* (Oxford: OUP): 279-298.
- Goldman, Alvin I. (2010), 'Epistemic Relativism and Reasonable Disagreement', in Richard Feldman and Ted Warfield (eds.), *Disagreement* (OUP): 187-215.
- Good, I.J. (1967), 'On the Principle of Total Evidence', *The British Journal for the Philosophy of Science* 17(4): 319-321.

- Hallsson, Bjørn G. (2019), 'The Epistemic Significance of Political Disagreement', *Philosophical Studies* 176: 2187-2202.
- Harsanyi, John C. (1985), 'Acceptance of Empirical Statements: A Bayesian Theory Without Cognitive Utilities', *Theory and Decision* 18 (1):1-30.
- Hazlett, Allan (2012), 'Higher-Order Epistemic Attitudes and Intellectual Humility', *Episteme* 9(3): 205-223.
- Holton, Richard (2004), 'Rational Resolve', *Philosophical Review* 113: 507-35.
- Holton, Richard (2014), 'Intention as a Model for Belief', in Manuel Vargas and Gideon Yaffe (eds.), *Rational and Social Agency: Essays on the Philosophy of Michael Bratman* (Oxford: Oxford University Press)
- Horowitz, Sophie (2014), 'Epistemic Akrasia', *Noûs* 48(4): 718-744.
- Jackson, Elizabeth (2019), 'Belief and Credence: Why the Attitude-Type Matters', *Philosophical Studies* 176: 2477-96.
- James, William (1896), 'The Will to Believe', reprinted (excerpt) in Reason and Responsibility, eds. Joel Feinberg and Russ Shafer-Landau (Wadsworth/Thomson 2005), pp 101-109.
- Jehle, David and Branden Fitelson (2009), 'What is the "Equal Weight View"?' *Episteme* 6(3): 280-93.
- Joyce, James (2005), 'How Probabilities Reflect Evidence', *Philosophical Perspectives* (19): 153-78.
- Kelly, Thomas (2005), 'The Epistemic Significance of Disagreement', *Oxford Studies in Epistemology* 1: 167-96.
- Kelly, Thomas (2008), 'Disagreement, Dogmatism, and Belief Polarization', *Journal of Philosophy* 105(10) 611-33.
- Kelly, Thomas (2010), 'Peer Disagreement and Higher Order Evidence', in Richard Feldman and Ted Warfield (eds.), *Disagreement* (OUP): 111-174.
- Kelly, Thomas (2013), 'Disagreement and the Burdens of Judgment', in David Christensen and Jennifer Lackey (eds.), *The Epistemology of Disagreement: New Essays* (OUP): 31-53.
- King, Nathan L. (2011), 'Disagreement: What's the Problem? Or A Good Peer is Hard to Find', *Philosophy and Phenomenological Research* 85(2): 249-72.
- Kitcher, Philip (1990), 'The division of cognitive labor', *Journal of Philosophy* 87(1), 5–22.
- Kornblith, Hilary (2010), 'Belief in the Fact of Controversy', in Richard Feldman and Ted Warfield (eds.), *Disagreement* (OUP): 29-52.
- Kraay, Klaus (2019), 'Elgin's community-oriented steadfastness', *Synthese*.
- Kraft, James (2007), 'Religious Disagreement, Externalism, and the Epistemology of Disagreement: Listening to Our Grandmothers', *Religious Studies* 43(4), 417–432.
- Kuhn, Thomas S. (1970). *The Structure of Scientific Revolutions*. Chicago: University of Chicago Press.

- Lackey, Jennifer (2010a), 'A Justificationist View of Disagreement's Epistemic Significance', in Adrian Haddock, Alan Millar, and Duncan Pritchard (eds.), *Social Epistemology* (Oxford: Oxford University Press, 2010): 298-325.
- Lackey, Jennifer (2010b), 'What Should We Do When We Disagree?', in Tamar Szabó Gendler and John Hawthorne (eds.), *Oxford Studies in Epistemology* 3 (Oxford: Oxford University Press, 2010): 274-93.
- Lackey, Jennifer (2013), 'Disagreement and Belief Dependence' in David Christensen and Jennifer Lackey (eds.), *The Epistemology of Disagreement: New Essays* (OUP): 243-68.
- Lackey, Jennifer (2014), 'Taking Religious Disagreement Seriously', in Laura Frances Callahan and Timothy O'Connor (eds.), *Religious Faith and Intellectual Virtue* (Oxford: OUP): 299-316.
- Lakatos, Imre (1970). 'Falsification and the Methodology of Scientific Research Programmes'. In Imre Lakatos and Alan Musgrave (eds.), *Criticism and the Growth of Knowledge*. Cambridge: Cambridge University Press, 91– 196.
- Lasonen-Aarnio, Maria (2014), 'Higher-Order Evidence and the Limits of Defeat', *Philosophy and Phenomenological Research* 88 (2): 314-345.
- Lasonen-Aarnio, Maria (2018), 'Enkrasia or evidentialism? Learning to love mismatch', *Philosophical Studies* 1-36.
- Lehrer, Keith, and Carl G. Wagner (1983), 'Probability Amalgamation and the Independence Issue: a Reply to Laddaga', *Synthese* 55(3): 339-346.
- Levi, Isaac (1967a), *Gambling with Truth: An Essay on Induction and the Aims of Science* (Cambridge: MIT Press).
- Levi, Isaac (1967b), 'Information and Inference', *Synthese* 17(4): 369-91.
- Levi, Isaac (1980), *The Enterprise of Knowledge: An Essay on Knowledge, Credal Probability, and Chance* (Cambridge, MA: MIT Press).
- Levi, Isaac (1991), *The Fixation of Belief and Its Undoing* (Cambridge, UK: Cambridge University Press).
- Lougheed, Kirk (2018a), 'The epistemic value of deep disagreements', *Informal Logic* 38(2): 263-92.
- Lougheed, Kirk (2018b), 'Is Religious Experience a Solution to the Problem of Religious Disagreement?', *Logos and Episteme*, 9(2): 173–97.
- McGrath, Sarah (2008), 'Moral Disagreement and Moral Expertise', *Oxford Studies in Metaethics* 3: 87-107.
- Muldoon, Ryan (2013), 'Diversity and the division of cognitive labor', *Philosophy Compass* 8(2): 117-25.
- Nissan-Rozen, Ittay and Levi Spectre (forthcoming), 'A Pragmatic Argument against Equal Weighting', *Synthese*.

- Owens, David (2002), 'Epistemic Akrasia', *The Monist* 85(3): 381-97.
- Pace, Michael (2011), 'The Epistemic Value of Moral Considerations: Justification, Moral Encroachment, and James' 'Will To Believe'', *Noûs* 45 (2):239-268.
- Palmira, Michele (2019), 'How to Solve the Puzzle of Peer Disagreement', *American Philosophical Quarterly*.
- Paul, Sarah K. (2015), 'Doxastic Self-Control', *American Philosophical Quarterly* 52 (2): 145-58 (2015)
- Pettit, Philip (2006), 'When to Defer to Majority Testimony—And When Not', *Analysis* 66(3): 179-87.
- Pittard, John (2014), 'Conciliationism and Religious Disagreement', in Michael Bergmann and Patrick Kain (eds.), *Challenges to Moral and Religious Belief: Disagreement and Evolution* (OUP).
- Pittard, John (2020), *Disagreement, Deference, and Religious Commitment* (New York, NY: Oxford University Press).
- Rosen, Gideon (2001), 'Nominalism, Naturalism, Epistemic Relativism', *Philosophical Perspectives* 15, *Metaphysics*: 69-91.
- Jeffrey Sanford Russell, John Hawthorne and Lara Buchak (2015), 'Groupthink', *Philosophical Studies* 172 (5): 1287-1309.
- Shogenji, Tomoji (2007) 'A Conundrum in Bayesian Epistemology of Disagreement.' Available online at: [http://fitelson.org/few/few\\_07/shogenji.pdf](http://fitelson.org/few/few_07/shogenji.pdf)
- Strevens, Michael (2003), 'The role of the priority rule in science', *Journal of Philosophy* 100(2): 55–79.
- Turnbull, Margaret Greta (forthcoming), 'Religious Disagreement is not Unique', in Matthew A. Benton & Jonathan L. Kvanvig (eds.), *Religious Disagreement and Pluralism* (Oxford: Oxford University Press).
- Turnbull, Margaret Greta and Eric Sampson (2020), 'How Rational Level-Splitting Beliefs Can Help You Respond to Moral Disagreement', in Michael Klenk (ed.), *Higher Order Evidence and Moral Epistemology* (New York: Routledge).
- Van Inwagen, Peter (1996), 'Is it Wrong Everywhere, Always, and for Anyone to Believe Anything on Insufficient Evidence?', Jeff Jordan and Daniel Howard-Snyder (eds.), *Faith, Freedom and Rationality* (Savage, Maryland: Rowman and Littlefield): 137-154.
- Van Inwagen, Peter (2010), 'We're Right. They're Wrong', in Richard Feldman and Ted Warfield (eds.), *Disagreement* (OUP): 10-28.
- Wedgwood, Ralph (2010), 'The Moral Evil Demons', in Richard Feldman and Ted Warfield (eds.), *Disagreement* (OUP): 216-46.
- Vavova Katia (2014), 'Moral Disagreement and Moral Skepticism', *Philosophical Perspectives* 28(1): 302-333.
- Weatherston, Brian (2005), 'Can We Do Without Pragmatic Encroachment?', *Philosophical Perspectives* 19: 417-43.