

# How to Endorse Conciliationism

Will Fleisher

Penultimate version, forthcoming in *Synthese*

## Abstract

I argue that recognizing a distinct doxastic attitude called *endorsement*, along with the epistemic norms governing it, solves the self-undermining problem for conciliationism about disagreement. I provide a novel account of how the self-undermining problem works by pointing out the auxiliary assumptions the objection relies on. These assumptions include commitment to certain epistemic principles linking belief in a theory to following prescriptions of that theory. I then argue that we have independent reason to recognize the attitude of endorsement. Endorsement is the attitude of resilient and committed advocacy which is appropriate for researchers to have toward their own theory. Recognizing the importance of endorsement, and of its resiliency, gives us reason to deny the epistemic principles that serve as auxiliary assumptions in the self-undermining objection. This defuses the objection, and provides additional support for the theory of endorsement.

## 1 Introduction

The epistemology of disagreement is concerned with the question of how you should respond, by the lights of epistemic rationality, to the discovery that you disagree with another person. Since disagreement is such a common occurrence both in inquiry and in everyday life, the problem of disagreement is at the heart of epistemology. The question is not about how you should respond to arguments or evidence offered by your interlocutor, but rather about how you should respond to the fact of the disagreement itself. The general view which suggests that (*ceteris paribus*) you should change your beliefs, merely on the basis of the discovery of the disagreement itself, is called *conciliationism*. Conciliationism, in both its more extreme and more moderate forms, has significant intuitive plausibility and is widely popular. However, it is beset by a difficult objection: the *self-undermining problem*.

In this paper, I will argue that we can solve this problem by appeal to the theory of *endorsement*. Endorsement is a propositional attitude that is distinct from both categorical (or full) belief and from degrees of belief, and which is governed by a distinct set of norms that I call inclusive epistemic rationality. Endorsement is the appropriate attitude of committed advocacy for researchers to have toward their own theories during inquiry.

Below, I will first offer some background on conciliationism, and then describe the two versions of the self-undermining problem. I will present a novel analysis of both versions that will help make clear what causes the problem. Then, I will offer a characterization of endorsement and the inclusive epistemic rationality which governs it. With this background in place, I will show how the theory of endorsement solves both versions of the problem.

## 2 Conciliationism

Consider the following version of a classic case from the epistemology of disagreement (Christensen 2007):

**Mental Math:** Suppose you go to lunch with your friend. When the check arrives, you decide to split it, and each do the math separately. You calculate that \$43 is the correct amount each of you should pay, and become quite confident in this claim. Your friend comes up with \$45, and is quite confident in their answer. You have no independent reason to think your friend is worse at elementary addition than you are. How should you respond?

Intuitively, many philosophers have thought that you should become less confident that the correct answer is \$43. After all, at least one of you has to be incorrect, and you have no reason to think that your friend is any less competent at basic arithmetic than you are. At least one of you has made a mistake, but for all you know the mistake is yours. Given this, it makes sense to be less confident in your own answer, where we understand this confidence in terms of degrees of belief (or credences).

More carefully, the intuition is that your degree of belief in the proposition “Half the check is \$43” should decrease, and your confidence in “Half the check is \$45 should increase.” Alternatively, for those who prefer talking in terms of categorical belief, the disagreement should cause you to give up your belief that \$43 is the correct answer, and should lead you to instead suspend judgment. Conciliationism generalizes from intuitions about cases like *Mental Math* to a general theory about the epistemically rational response to disagreement.

Conciliationist theories are those which (at least often) call for a subject to significantly change her belief state in response to the mere fact of disagreement. That is, the discovery of the disagreement itself is the evidence (or at least the impetus) that leads to the change in belief. More specifically, conciliationist theories call for a subject to change her beliefs to be significantly closer to those of her disagreeing interlocutor, at least in many cases of disagreement.

As I will use the term in this paper, a conciliationist view is one that will agree with the following general principle, or a principle similar to it:

**Conciliationism (CV):** If a subject has a certain credence  $c_1$  with respect to  $p$ , and then learns an epistemic peer has a different cre-

denance  $c_2$  towards  $p$ , then the subject should (*ceteris paribus*) adopt a new credence  $c_3$ , which is significantly closer to  $c_2$  (than  $c_1$  is).<sup>1</sup>

Several aspects of the above principle need explanation. First, as I will understand things, the kind of disagreement mentioned in CV consists in learning that a peer has a different credence towards a relevant proposition.<sup>2</sup> Second, an epistemic peer is another person whose epistemic position is just as good as the subject's. This notion of epistemic position encompasses both the available evidence and competence at evaluating evidence, relative to a subject matter. That is to say, a genuine peer has all the same evidence as the subject, and is just as good at evaluating evidence as the subject. Intuitively, in the mental math case, you and your friend are peers because you both have access to the check and a similar mastery of elementary arithmetic. Much ink has been spilled about the nature of epistemic peerhood, but this debate does not concern us here.<sup>3</sup> The self-undermining problem arises for conciliationism on any of the plausible ways of cashing out the notion of an epistemic peer. Moreover, the problem arises for any view that recommends conciliation in cases of non-peer disagreement, e.g., in cases where the disagreeing interlocutor is in an epistemic superior (or even inferior) position. Most conciliationist views will call for conciliation in these cases as well. For our purposes, then, "epistemic peer" can be read as any interlocutor with at least an approximately similar epistemic position.

Third, CV concerns cases where a subject discovers actual disagreement. Some have argued that merely potential disagreement should be treated the same, epistemically, as actual disagreement.<sup>4</sup> I think that conciliationism is only plausible for actual disagreement. More to the point, the self-undermining objection involves actual disagreement among epistemologists. So, that will be our focus here.

Conciliationism is a category of views that contains a variety of more specific theories which specify how a subject should change her beliefs, and by how much. The most famous of these is the *equal weight view* (EWV) (Christensen 2007; Elga 2007). According to EWV, a subject is always required to significantly conciliate in circumstances of peer disagreement. Specifically, the subject must give their peer's judgment equal weight to their own (provided the subject does not gain evidence independent of the disagreement that the interlocutor is not really a peer).<sup>5</sup>

<sup>1</sup>This version of the generalized CV principle is adapted from Elga (2010) and Matheson (2015a). I have framed CV in terms of credences, but a similar principle can be formulated for categorical or full beliefs. For ease of exposition, I will focus on the credence version of the principle, but my proposed solution will apply to both versions.

<sup>2</sup>Here I am following Lasonen-Aarnio (2013)

<sup>3</sup>For an overview of this debate, and some more nuance about the available positions, see King (2012) and Frances and Matheson (2018).

<sup>4</sup>For an overview of this dispute, see Matheson (2015c).

<sup>5</sup>It turns out that a precise specification of "giving equal weight" is hard to formulate, especially if one wants to maintain compatibility with Bayesianism. See Gardiner (2014); Jehle and Fitelson (2009); Lasonen-Aarnio (2013); Rosenkranz and Schulz (2015).

EWV has often been seen as an extreme view, as it calls for significant revision of credence in all cases of peer disagreement, and puts stringent independence requirements on dismissing an interlocutor's peerhood. Several philosophers have proposed more "moderate" views. These are meant to require less significant conciliation, and allow for "downgrading" purported peers on the basis of the very disagreement in question.<sup>6</sup> Moderately conciliatory views are committed to CV, as I have construed it. This classification is helpful because both extreme and moderate versions of conciliationism face the self-undermining problem.

Conciliationism's opponent is often called steadfastness, or the steadfast view (SV). Steadfast views claim that one should respond to disagreement by maintaining one's initial credence. That is, (in at least many cases) the mere fact of disagreement does not provide any reason to lower one's confidence or change one's belief state. In general, steadfast views will accept the following principle:

**Steadfastism (SV):** If a subject has a certain credence  $c_1$  with respect to  $p$ , then (*ceteris paribus*) the subject should maintain  $c_1$  when she learns an epistemic peer has a different credence  $c_2$  towards  $p$ .<sup>7</sup>

There are also a number of different steadfast views, but their differences will not concern us.<sup>8</sup>

The self-undermining problem is an objection to conciliationism in general, in favor of steadfast views in general. It is supposed to show that even those with conciliationist intuitions in cases like *Mental Math* should nonetheless give up conciliationist theories. My project here is to show that this objection fails once we adopt the theory of endorsement. Some philosophers do not seem to share the conciliating intuitions from the mental math case, and some are motivated to be steadfasters for other reasons. The arguments here are not meant to persuade such philosophers to give up the steadfast view. Rather, they are simply meant to block a significant objection to conciliationism.

The kind of disagreement we see in the mental math case is often taken to work as a kind of *higher-order defeater* (Christensen 2010a; Lasonen-Aarnio 2010, 2014; Schoenfield 2018; Titelbaum 2015). The evidence provided by the

<sup>6</sup>For examples of moderate conciliatory views, see Kelly (2010), Sosa (2010), Lackey (2008), Levinstein (2015), Wiland (2018), and Worsnip (2014). The purpose of the "significantly" qualifier in my version of CV is meant to avoid issues dealing with a moderate view so weak that it meets the letter of CV, but fails to call for any serious revision of belief. Such a theory might avoid the problem, but would fail to be an interesting or plausible view.

<sup>7</sup>Note that I am calling the principle an "-ism" because I want to highlight that it involves a theoretical commitment to such a principle, as that is what leads to the self-undermining problem. Also note that, as stated, it might be possible to build a view that packed enough into the *ceteris paribus* clauses of both CV and SV to technically satisfy both. But the resulting view would either a) still face the self-undermining objection, or b) have to include an ad hoc provision against it. Moreover, I don't think anyone actually defends a view that would be well-described in this way.

<sup>8</sup>For versions of steadfast views, see Kelly (2005), Titelbaum (2015), and Lasonen-Aarnio (2010, 2014). For additional background on the peer disagreement literature, see Christensen (2009), Matheson (2015c), and Frances and Matheson (2018).

disagreement in cases like mental math is evidence that the subject has made a mistake. This is evidence that something is wrong with the subject's beliefs, not evidence directly against the proposition in question (e.g., that half the check is \$43). Disagreement is not the only kind of higher-order defeater. For instance, finding out that I might be suffering from hypoxia (oxygen deprivation) can give me reason to doubt my belief-forming methods (Christensen 2010b). In this paper I will focus exclusively on disagreement. However, the solution offered may also help in resolving similar difficulties for other kinds of higher-order defeaters.

### 3 The Self-Undermining Problem

The self-undermining problem is perhaps the most pressing difficulty for conciliationism. Epistemologists have developed this objection in a number of ways, but since the solution I will offer for this problem is applicable to any of them, I will focus on two particular ways in which the problem arises.

The basic idea of the self-undermining problem is that conciliationism, when applied to the epistemology of disagreement, recommends giving up (or significantly lowering credence in) conciliationism. That is, advocating conciliationism by taking part in the epistemology of disagreement gives you strong reason, by conciliationism's own lights, to give up conciliationism.

Here is a case we can use to illustrate the problem:

**Connie the Conciliator:** Connie is an epistemologist. After reading the early literature on disagreement, she is attracted to conciliationism. She finds herself becoming a committed advocate of the view, exploring and defending it. Later, she is confronted with disagreement from some of her peers, who are steadfasters. As a conciliationist, she dutifully lowers her confidence in conciliationism, in light of this disagreement. She now has less than .5 credence in the theory, and certainly less credence than would be required to rationally believe the theory.<sup>9</sup>

There are two main versions of the self-undermining problem that have been proposed. The first version, which I will call *self-undermining justification*, is due originally to Plantinga (1999). The second version, which I will call *self-undermining inconsistency*, was proposed by Elga (2010) and Weatherson (2013). We can use Connie's case to illustrate each version.

#### 3.1 Two versions of self-undermining

The basic idea of the self-undermining justification problem is that adopting conciliationism causes a subject to lose her justification to believe conciliationism. The term "adopting" is meant to be a neutral term for being committed to

<sup>9</sup>Compare this example to cases from Elga (2010) and Weatherson (2013).

a theory.<sup>10</sup> This could either mean believing the theory, or as I will argue later, endorsing it. The self-undermining justification problem arises when we interpret adopting to mean believing. Connie's case illustrates how the problem works by showing what will happen if we suppose someone in circumstances like our own follows CV when there is disagreement about CV. We can use this hypothetical case to give the argument:

### Self-undermining justification argument

1. Connie adopts CV, and then meets several disagreeing peers.
2. CV (given this disagreement) recommends lowering her credence in CV below what is required for justified belief.
3. Connie should follow this recommendation.
4. If she does so, she is no longer justified in believing CV.
5. If Connie is not justified in believing CV, she should give it up (no longer be committed to it).
6. Thus, Connie should give up CV.<sup>11</sup>

Premise one is the assumption that Connie becomes committed to conciliationism, where usually this is interpreted as believing it. The first conjunct is necessary because Connie will only follow CV if she adopts it. CV being true without her knowing about it won't generate the problem. Premise two states that after Connie confronts the disagreement with her peers, conciliationism recommends lowering her confidence in CV. Because she is committed to conciliationism, according to premise three Connie should follow this recommendation. Because her confidence is too low in the theory, she is no longer justified in believing it (premise 4). If you are no longer justified in believing a theory, you should give it up, i.e., no longer be committed to it (premise 5). So Connie should not be committed to conciliationism.

If this argument is sound, this is a serious problem for conciliationism. It looks like adopting the theory is unstable: adopting it, in the current circumstances of disagreement, seems to lead one to be rationally required to give up the view. In other words, adopting the view is unstable rationally speaking. This looks bad for conciliationism.

The self-undermining inconsistency objection is perhaps even worse: it alleges that conciliationism is incoherent because it gives inconsistent advice in many circumstances in epistemology of disagreement. This can also be illustrated by appeal to the argument as it applies to Connie's case. Suppose that Connie proceeds as before, and has met several peers who disagree with her.

### Self-undermining inconsistency argument

1. Connie adopts CV, and then meets several disagreeing peers.

<sup>10</sup>This language is inspired by Elga (2010), who talks in terms of adopting and giving up views. Weatherson (2013) gives his version of the objection in terms of both belief and in terms of "theories to trust."

<sup>11</sup>See Plantinga (1999), Weatherson (2013), Decker (2014) and Matheson (2015a) for this version.

2. CV (given this disagreement) recommends lowering her credence in CV below her credence in SV.
3. If Connie has greater confidence in SV than CV, then she should follow the rules prescribed by SV.
4. Thus, CV recommends following the rules prescribed by SV.
5. CV recommends both lowered credence in CV, and following SV (which recommends maintaining her initial credence).
6. Thus, CV gives inconsistent advice.
7. One should not maintain a commitment to a theory which gives inconsistent advice.
8. Therefore, Connie should give up CV.<sup>12</sup>

Premise one's first conjunct is again necessary so that Connie will follow the rule, which generates the purported inconsistency. Premise two here suggests that, in Connie's case, there is adequate disagreement to warrant Connie actually having less than 1/2 confidence in CV, and so having greater confidence in SV. This is a reasonably plausible situation for Connie to find herself in, given the current circumstances in the epistemology of disagreement. Premise three is plausible if what one should do depends on the theory one has most confidence in (something I will take issue with shortly). Premise five follows from the different commitments in SV and CV. Steadfastism recommends something inconsistent with conciliationism in this case: SV recommends that Connie maintain her level of confidence in CV. CV, meanwhile recommends that Connie lower her confidence in CV. It's impossible for her to follow both recommendations. Since CV gives inconsistent advice that cannot be followed, Connie should give it up as a theory of what rationality requires her to do. One cannot be rationally required to do something impossible.

If this argument is sound, then it at least appears that conciliationism is self-undermining, because adopting the theory results in receiving inconsistent advice. This looks even worse for the view than the first version of the problem. Not only does adopting the view leave one unjustified, it turns out that adopting the view is incoherent.<sup>13</sup>

### 3.2 The nature of the problems

There are two things to note about these arguments. The first is that self-undermining arises in the context of doing epistemology. The issue is about what we should do when we disagree about *conciliationism*, a *theory* about the correct response to disagreement. Anyone who engages with this issue is thereby doing epistemology. This is the reason for talk of adopting a *theory*:

<sup>12</sup>See Elga (2010) and Weatherson (2013) for this version, which is inspired in part by Lewis (1971).

<sup>13</sup>For some related problems for conciliationism, see Mulligan (2015). Mulligan's paradoxes involve problems where conciliationism offers conflicting advice. However, these paradoxes don't result from applying conciliationism to disagreement about CV, but instead arise in cases of disagreement about peerhood. Thus, I think they are distinct from the self-undermining objections discussed here, and require separate treatment.

the subjects in question are engaged in epistemological theorizing. What attitude should be involved in this adoption of a theory is one of the topics of the present paper. But the self-undermining objection only arises for someone who is committed to the theory of conciliationism. Someone who is prone to conciliate, but never thinks about it, will not have their beliefs (or anything else) undermined in this way. They may hold no attitude whatsoever to the theory.

The second thing to note is the epistemic nature of the objection. Neither version of the problem amounts to a claim that the conciliatory view is self-refuting or necessarily irrational to believe. The state of affairs where someone adopts (or believes) conciliationism is not incompatible with the truth of conciliationism. Both versions of the self-undermining problem are epistemic problems: one cannot rationally adopt (believe) the theory, because doing so undermines the theory's justification, or results in the subject getting inconsistent advice. This is not the same thing as self-refutation, which occurs when believing a claim (or adopting a theory) ensures the falsity of that claim. Self-refutation is illustrated by the Liar sentence, "This sentence is false." Alternatively, consider the claim "No one has any beliefs"; belief in that claim is self-refuting.<sup>14</sup>

Moreover, the self-undermining problem does not arise with necessity. Even regarding the inconsistency version, the objection is not that conciliationism always and necessarily provides incoherent advice. Rather, the problems arise from the existence of actual disagreement with peers or superiors. But such disagreement is a contingent feature of present circumstances. If, as conciliationists hope, CV some day becomes the consensus view, then the disagreement will cease. If that occurs, then believing conciliationism would not undermine justification in conciliationism, and CV would give consistent advice. Contrast this with something like the Moorean proposition "It is raining, but I do not believe it is." One cannot ever be in a position to rationally believe such a claim, no matter what happens (Sorensen 1988). Moorean propositions face an epistemic problem that arises necessarily in all circumstances.

This is not true for the self-undermining problem. It is a problem of the here and now. It would no longer be a problem if there were no longer disagreement about CV.<sup>15</sup> What conciliationists need is a solution to self-undermining that buys them time: time in which to rationally pursue the theory despite disagreement about it. This would allow them to continue their pursuit of the theory so that new arguments may be presented, new evidence collected, and so that a new consensus might emerge. The solution I propose below, by appeal to endorsement, is meant to offer that opportunity.

A variety of solutions have been proposed for the self-undermining problem, in both its versions. After raising both problems, Elga (2010) suggests that CV must be amended so that it does not apply to itself. This is supposed to be

---

<sup>14</sup>For more on the differences between self-undermining and self-refutation, see Matheson (2015a) and Decker (2014).

<sup>15</sup>Again, this is a point made by Matheson (2015b).



acceptable because the same must be true for all inductive methods (Lewis 1971). Bogardus (2009) offers a different solution. He suggests that we must distinguish between knowledge from testimony and knowledge by acquaintance, and that we know rational principles like CV through acquaintance. Pittard (2015) suggests that remaining steadfast in CV actually respects the spirit of CV better than giving it up, as remaining steadfast is what the disagreeing interlocutor is advocating.<sup>16</sup>

Each of these views has associated worries. Concerning Elga's solution, one might worry that the exemption seems an ad hoc way to avoid the objection (Weintraub 2013). About Bogardus' solution, one might wonder whether a topic of deep philosophical disagreement was ever knowable by acquaintance, and if it was, whether that knowledge should survive the available defeaters. Regarding Pittard's solution, the concern is that there are multiple ways to respect the spirit of conciliationism, and we need more reason to prefer remaining steadfast. There is not space here to adequately evaluate these other views and their purported problems. Instead, I will propose an alternative solution that is independently motivated by the work it does in the epistemology of inquiry.

The solution that I propose requires that we recognize that both versions of the problem rely on additional commitments, over and above a commitment to conciliationism. First, each requires that we treat "adopting" a theory as believing it. As I will argue later, this assumption should be abandoned. Second, and more importantly, each of these arguments requires a background commitment to a general epistemic principle.

The self-undermining justification objection relies on the following principle:

**Unjustified Theory:** If a subject *S* is not justified in believing (or would not be justified in believing) a theory, then *S* is rationally required to give up (not be committed to) that theory.

This principle is a general expression of the principle in premise 5 of the self-undermining justification problem. Without this principle, the conclusion does not follow. I will argue that although this principle seems intuitive, it is in fact false. Being committed to a theory is rationally compatible with being unjustified in believing the theory. This is because the proper way to be committed to a theory during inquiry is to endorse it.

The self-undermining inconsistency problem relies on a different epistemic principle:

**Most Credence:** If a subject has the most credence in a theory, then the subject should follow any rules prescribed by that theory.<sup>17</sup>

<sup>16</sup>For an overview of related literature, see Frances and Matheson (2018). Wiland (2018) and Matheson (2015b) both offer solutions that are similarly sensitive to the fact that self-undermining requires additional background commitments, though their solutions are differently motivated.

<sup>17</sup>Note that *Most Credence* could also be formulated in more general way, so that it concerns

This is a special kind of *enkratic principle*. It connects the subject's credence in competing epistemic theories with whether they should act on the prescriptions contained in those theories.<sup>18</sup> In particular, it requires subjects to follow prescriptions of the theory have the highest confidence in.

The *Most Credence Principle* underwrites the third premise in the inconsistency argument. Premise three does not follow without a commitment to an enkratic principle of this kind. This is because conciliationism is a theory that only prescribes how subjects should change their beliefs in light of disagreement. Conciliationism itself does not include any advice about when to follow rules prescribed by theories one has high credence in.<sup>19</sup> CV and Connie's evidence imply that she should lower her confidence in CV and raise it in SV. CV itself says nothing about how to act upon rules contained in theories like SV. It is only in combination with the *Most Credence Principle* that Connie's newfound high credence suggests that she should now follow the dictates of steadfastism. The self-undermining inconsistency argument is only valid with this suppressed premise. Without *Most Credence*, there is no inconsistency.<sup>20</sup>

Each version of the self-undermining argument relies on a particular epistemic principle. I will argue that each of these principles is false. One can adopt a theory even if one would be unjustified in believing it. One should not always follow the rule that one has most credence in. In order to motivate these claims, I will present a general theory about inquiry which denies them. According to this theory, the appropriate kind of commitment to a theory is an attitude called endorsement. In the next section, I will present this theory of endorsement, before returning to argue that the theory provides strong motivation for denying the two principles.

---

belief in any normative claim or any proposition which includes rules about how to act. Here, I formulate it in terms of theories because self-undermining arises when evaluating the theory of conciliationism.

<sup>18</sup>Generally, enkratic principles constrain what kinds of belief states and actions (or intentions) are rationally compatible. They are so named because they rule out the rationality of *akrasia*: believing that one should do something, and yet not doing it. Enkratic principles in epistemology usually involve prohibiting akratic beliefs: believing something (to a certain degree), while (justifiedly) believing that one should not so believe. For more on enkratic principles, see Horowitz (2014) and Broome (2013).

<sup>19</sup>Lasonen-Aarnio (2014) and Pittard (2015) both briefly mention this point, but neither explores it in depth.

<sup>20</sup>Weatherson (2013) and Elga (2010) (in a footnote) both suggest a third formulation of the self-undermining problem. This alternative version suggests that CV is self-undermining if it suggests *any* lowering of credence in itself. The argument actually requires commitment to a different enkratic principle, which we can call the *weighted average principle*. This principle is that one should follow rules which are derived from a weighted average of rules contained in all the theories one has some credence in (with the weights provided by the credences). The same endorsement solution applies to this version of the self-undermining problem as the one appealing to the *Most Credence Principle*. So, I will focus on the version relying on *Most Credence*, for ease of exposition.

## 4 Endorsement

Researchers often act as committed advocates for their theories. Such inquirers advocate for their theories, assert their superiority, and defend them doggedly against difficult objections. Moreover, that researchers act this way is a good thing: such practices are constitutive of healthy inquiry. But in ordinary contexts, such commitment and advocacy is explained by the fact that people believe the claims in question. However, belief and even high credence in one's own theory seems irrational given the systematic and pervasive disagreement researchers confront in cutting-edge inquiry. Moreover, there are additional reasons to doubt that belief is warranted. In cutting-edge fields characterized by disagreement, the available evidence underdetermines which theory to choose. Moreover, most theories which have been proposed have not ultimately been accepted, but have been discarded as false (this is a version of the pessimistic meta-induction (Psillos 1999)).

Despite the usual connection between belief and committed advocacy, and despite the apparent irrationality of belief in the relevant contexts, we do not want our theory of epistemic rationality to count researchers as irrational for being committed advocates of their views. Nor do we want it to prohibit committed advocacy, which is beneficial to inquiry. There is thus a tension between our theory of individual epistemic rationality, and the goals of collective inquiry.

In order to resolve this apparent tension, I propose that we recognize a distinct doxastic attitude, one which is appropriately governed by rational norms concerning collective inquiry. I call this attitude *endorsement*, and the norms which govern it *inclusive epistemic rationality* (Fleisher 2018, 2019). As I will argue below, this theory of endorsement provides a solution to both versions of self-undermining. In the rest of this section, I will first provide a characterization of endorsement, followed by a brief explanation of inclusive epistemic rationality.

### 4.1 The Nature of Endorsement

Endorsement is a propositional, doxastic attitude. It embodies the resilient commitment and advocacy that researchers should have toward their theories during inquiry. The following characterization is a part of the attitude's functional profile that we can use to distinguish it:

**Endorsement:** Endorsement is a propositional attitude. It is an attitude of resilient commitment and advocacy. *S* endorses *p* in a research domain *d* only if:

1. *S* is disposed to assert that *p*, or otherwise express commitment to *p* (in *d*).
2. *S* takes herself to be obligated to defend *p* (in *d*).
3. *S* treats *p* as a premise in her further reasoning (in *d*).
4. *S* shapes her research program in *d* (in part) based on *p*.

5. *S* is resiliently committed to *p* (in *d*).
6. *S* takes *p* to be a live option, i.e., she does not know *p* is false.
7. In endorsing *p*, *S* aims to promote successful inquiry.<sup>21</sup>

This characterization is not a set of necessary and sufficient conditions. Instead, it simply provides some characteristic markers as a way of distinguishing the attitude from other propositional attitudes (e.g., belief, desire, or hope).<sup>22</sup> There are a few things that need to be clarified in this characterization. First, a subject endorses a proposition within a particular research domain. This makes endorsement a “fragmented” attitude, meaning it is compartmentalized to one part of the subject’s mental life, rather than being a global feature of their mental state.<sup>23</sup> This fragmentation helps explain the way subjects can behave as committed advocates within a research domain, but not be willing to accept the theory (or gamble on it for high stakes) outside of the domain.

Endorsement is related to other kinds of doxastic attitudes, such as categorical belief and credence. It is an attitude toward the truth of a statement or proposition. However, it differs from each of these other attitudes in a variety of ways. The last three of the necessary conditions in particular allow for the attitude to be distinguished from categorical belief. These three conditions require dispositions to act in ways that are not required for a belief. One need not be obligated to engage in a research program on the basis of any belief one has. Endorsing a theory in the context of inquiry, however, involves a practical commitment to engage in such research.

The resilient commitment requirement in condition (5) also serves to distinguish endorsement from categorical belief. The sense of resiliency here involves a maintenance of the commitment in the face of contrary evidence. In this sense, belief is less resilient than endorsement. When faced with strong contrary evidence, such as a purported counter-example, a rational subject should give up the belief. However, a researcher’s endorsement of a theory often survives the discovery of significant contrary evidence. As I will suggest in the next sections, this resiliency in the face of contrary evidence is rational, and beneficial for collective inquiry.

Endorsing *P* is compatible with suspending judgment (also called “withholding belief”) about whether *P*. Generally, endorsing a theory will involve suspension of judgment on the question that the theory provides an answer to.<sup>24</sup> Importantly, endorsing *P* is compatible with having a particular credence

<sup>21</sup>There are many theories about successful inquiry compatible with the endorsement framework. I prefer veritism (Goldman 1987; Pettigrew 2016).

<sup>22</sup>This notion of endorsement is inspired by the acceptance/belief distinction, especially by the work of L.J. Cohen (1989), Levi (1980), Maher (1993), Whitt (1990) and Van Fraassen (1980). Recently, several philosophers have recognized the need for a provisional acceptance attitude of some kind, e.g., Goldberg (2013), McKaughan (2007), and Barnett (2019). Elgin (2010) appeals to Cohen’s notion of acceptance to help with disagreement problems. I think endorsement does the best job of playing this provisional acceptance role.

<sup>23</sup>For more on fragmentation, see Egan (2008); Elga and Rayo (2019); Lewis (1982); Rayo (2013); Stalnaker (1984).

<sup>24</sup>For this notion of suspension, see Friedman (2017).

in  $P$ . On my account of rational endorsement, what it is rational to endorse will in part depend on one's credences. I think it is possible to both endorse and believe a proposition. However, it will often be rational to endorse something it is irrational to believe, for reasons to be discussed below.

Endorsement is the appropriate attitude for researchers to take toward the theories they pursue as committed advocates. It is a provisional attitude, one researchers take toward a theory when they think doing so will best promote successful inquiry. Categorical belief, on the other hand, should be taken only toward answers to settled questions (or at the very least, toward answers with a very high degree of evidence in their favor).<sup>25</sup> Because of this difference in roles, belief and endorsement are governed by different epistemic standards. For instance, one cannot rationally (or justifiedly) believe a proposition  $P$  when one takes  $\neg P$  to more likely: in other words, you should not believe  $P$  if your credence in  $P$  is less than .5. Moreover, you should not believe  $P$  when you take a competitor theory  $Q$  to be more likely to be true, i.e., when your credences have it that  $Pr(P) < Pr(Q)$ . On my account, endorsement can be rational in both of these circumstances. This is due to its nature as a provisional attitude taken during inquiry.<sup>26</sup>

## 4.2 Inclusive Epistemic Rationality

With an account of the attitude on the table, we can turn now to a characterization of the norms that govern it. Part of what distinguishes endorsement from belief and credence is the distinctive kind of epistemic reasons to which endorsement is sensitive. I call these *extrinsic epistemic reasons*, and the normative framework that includes such reasons *inclusive epistemic rationality*. Sensitivity to extrinsic epistemic reasons is an important aspect of the theory of endorsement's solution to the self-undermining problem.

Inclusive epistemic rationality is inclusive of both intrinsic and extrinsic epistemic reasons.<sup>27</sup> Intrinsic epistemic reasons are reasons which are about, or indicate, the truth of the proposition in question. They are reasons to think that a proposition is true. If  $Q$  is evidence for  $P$ , then  $Q$  is an intrinsic epistemic reason (for  $P$ ). Other intrinsic reasons are things which serve as necessary con-

<sup>25</sup>For the idea of belief as a settling attitude, see Friedman (2017) and Staffel (2019).

<sup>26</sup>These features help distinguish the view from competitors, e.g., those proposed by Goldberg (2013) and Barnett (2019). For these reasons, only endorsement has the resources to give the solution to self-undermining on offer below. Goldberg's account requires that one be more confident than not before it is appropriate to champion the view. Barnett's account of inclination requires that the subject have most confidence in a theory. Neither appeals to extrinsic epistemic reasons.

<sup>27</sup>This intrinsic/extrinsic terminology is originally due to Steel (2010), though I develop the distinction differently than he does. Loughheed and Simpson (2017) are concerned with a similar distinction. On my view, extrinsic epistemic reasons include considerations of whether a theory or research program is worthy of pursuit in the sense suggested by L. Laudan (1978). Thus, a number of philosophers have explored what I take to be extrinsic epistemic reasons. For an overview of the literature on pursuitworthiness, see McKaughan (2007) and Whitt (1990). Relevant work on this topic includes R. Laudan (1987); Longino (1990); McKaughan (2008); McMullin (1976); Nickles (1981); Šešelja, Kosolovsky, and Straßer (2012); Šešelja and Straßer (2013, 2014); Solomon (1994); Whitt (1992). For more about my take on the distinction, see Fleisher (2018).

ditions for truth: e.g., that a theory is consistent is (some) reason to think it is true, and therefore is an intrinsic reason. Intrinsic epistemic reasons indicate, or point to, the truth of the proposition in question.<sup>28</sup>

Extrinsic epistemic reasons are reasons concerning what will promote healthy inquiry. They are reasons about what will lead to getting more truths, or more knowledge, in the long run. Such reasons might concern only the productivity of an individual engaged in research alone. However, extrinsic epistemic reasons can also be reasons concerning the promotion of the goals of collective inquiry.

Extrinsic epistemic reasons go beyond reasons to think a theory is true. They are reasons in favor of a proposition that concern how taking some attitude (or action) regarding the proposition will affect inquiry. For instance, the testability of a theory is often taken to be an important feature for the purposes of scientific inquiry. This is not because testability indicates truth (most testable theories have turned out false) (Steel 2010). What testable theories have going for them is their suitability to inquiry. Because of this, that a theory is testable is an extrinsic epistemic reason to pursue it. Other extrinsic epistemic reasons concern the division of labor in collective inquiry: if a theory (or paradigm, or project) has very few people working on it, this is an extrinsic epistemic reason to endorse and pursue it. This is because collective inquiry is likely to go better if there is a better distribution of cognitive labor.<sup>29</sup>

On my account, extrinsic epistemic reasons are not merely pragmatic reasons. They deserve to be considered “epistemic” because they are about inquiry and the pursuit of knowledge. But they are not the kinds of reasons that have been typically treated as epistemic, as they do not indicate the truth of the proposition they relate to. Nonetheless, I think it is worth distinguishing between reasons concerning the promotion of practical ends (like becoming rich and famous) and reasons which concern successful inquiry. Extrinsic epistemic reasons are not about merely practical benefits, but about long-term or collective epistemic benefits. For instance, if a researcher chooses to endorse a theory because it has too few defenders, this is a reason that has to do with promoting healthy inquiry, not with fame or fortune. It is useful, theoretically, to treat these reasons as distinct from practical reasons. An account of epistemic rationality that leaves these reasons out, such as an evidentialist account (Shah & Velleman 2005), leaves out a great deal of our seemingly epistemic activity. Meanwhile, a fully pragmatist view like Rinard’s (2015), that denies the existence of epistemic normativity as a real kind, misses important distinctions between different kinds of reasons. For instance, it does not distinguish between reasons concerning fame and fortune, on the one hand, and reasons concerning the success of collective inquiry, on the other. That *working on theory A will lead to a better distribution of cognitive labor* seems like an epistemic reason, whereas that *working on theory A will make me rich* does not.

<sup>28</sup>I am using the normative language of reasons, but this is largely for convenience. The framework presented here is compatible with using “ought” language, or “value” language.

<sup>29</sup>See Kitcher (1990) and Strevens (2003).

Of course, researchers may (and most likely will) also be motivated by pragmatic considerations. However, it seems a common occurrence that researchers are motivated by a desire to contribute to healthy inquiry. Researchers who are thus motivated should not be evaluated as irrational or unjustified by our best theory of epistemic rationality. Rather, our theory should vindicate and encourage researchers who are sensitive to reasons concerning the overall health of collective inquiry. Thus, we need a category of genuinely epistemic reasons which go beyond reasons to think a theory is true. The category of extrinsic epistemic reason plays this role.

One distinctive feature of endorsement is that it is sensitive to both intrinsic and extrinsic epistemic reasons. This is distinctive because it is highly plausible that belief (both categorical and partial) is not sensitive to extrinsic epistemic reasons, but only to intrinsic ones. This is illustrated by the common intuition that it is epistemically irrational to believe in a way that does not fit one's evidence, or in a way that is known to be unreliable. Beliefs should fit the evidence, or be reliably formed (preferably both). The rational insensitivity of beliefs to extrinsic epistemic reasons is also illustrated by the common intuitions to epistemic bribery cases.<sup>30</sup> It looks generally inappropriate to take on a belief that is likely to be false, or is not supported by one's evidence, in order to gain other true beliefs. If one is convinced of atheism on the evidence, it is epistemically irrational to believe in God, even if that belief will provide one with a great deal of research funding and thus many more true beliefs in the long run. This is an extrinsic reason to be a theist, and thus does not support belief in theism.

Endorsement, on the other hand, is sensitive to reasons having to do with healthy inquiry, and is thus an appropriate attitude to take toward theories which are promising, or have too few defenders, but which the subject does not take to be highly likely to be true. What it is rational for a researcher to endorse depends on the balance of intrinsic and extrinsic epistemic reasons: that is, on inclusive epistemic rationality. Sometimes, the extrinsic epistemic reasons to endorse a theory will outweigh the intrinsic epistemic reasons in favor of its competitors, and it will be epistemically rational for a researcher to endorse a theory which she takes to be less probable than its competitors. Thus, there will be at least some cases of appropriate endorsement which would be inappropriate epistemic bribery were the subject to instead believe.

The sensitivity of endorsement to extrinsic epistemic reasons also helps to explain the resilience of endorsement in the face of contrary evidence. A researcher who endorses a theory that is faced with a difficult objection, or purported counterexample, might rationally lower her credence significantly in the theory. Still, her extrinsic epistemic reasons may outweigh this loss of credence, and make it rational for her to maintain her endorsement. For instance, she might be one of very few researchers working on the theory, or it might be a significant cost to change research programs (based on available laboratory space or research equipment).

<sup>30</sup>For these cases see Firth (1981), Jenkins (2007), Berker (2013), and Greaves (2013).

Rational sensitivity to extrinsic epistemic reasons helps the practice of endorsement promote healthy inquiry. Researchers will have reasons to endorse different theories from one another, even when in agreement about the available evidence. This will promote a useful distribution of cognitive labor, and will also lead to beneficial disagreement that helps researchers reason more effectively, and which promotes better evidence collection.<sup>31</sup> Sensitivity to extrinsic reasons will also help promote resilient commitment, as extrinsic reasons to maintain a commitment will remain even in the face of difficult contrary evidence. Thus, endorsement and its inclusive epistemic rationality provide significant benefits to inquiry.

I think the theory of endorsement, comprised by the account of the attitude and of inclusive epistemic rationality, is an intuitively plausible way of understanding and vindicating the committed advocacy of researchers. The primary justification for the theory is based on its explanatory and normative benefits. There are two main parts to this justification. The first is an inference to the best explanation: the theory best explains our intuitive judgments in a variety of cases. It also solves several problems in social epistemology and general philosophy of science. The second justification is related but concerns normative benefits. The thought is that a community that includes endorsers who are sensitive to inclusive epistemic rationality will inquire more successfully. Part of the argument for endorsement is the solution the theory provides for self-undermining. I have provided additional parts of this argument elsewhere (Fleisher 2018, 2019).

## 5 The Endorsement Solution to Self-undermining

As I argued above, the two versions of the self-undermining problem for conciliationism require appeal to additional assumptions beyond commitment to CV. The endorsement solution to self-undermining works by providing strong, independent motivation for denying these additional commitments.

The first of these assumptions is that adopting or committing to a theory requires belief in that theory. I have argued, instead, that endorsement is the appropriate attitude of commitment to a theory. So, being a conciliationist means having an attitude of endorsement toward CV (not belief). This claim is meant to vindicate current conciliationists by appeal to the theory of endorsement, which has independent motivation and application beyond the epistemology of disagreement. I think it accurately describes many conciliationists. Moreover, for those who really do believe CV, it offers an alternative kind of commitment that fits better with the content of CV. Thus, the first assumption needed for self-undermining is false.

The second assumption required for each version of self-undermining was the truth of certain epistemic principles. I will now argue that these specific principles are false.

---

<sup>31</sup>For some of the empirical evidence supporting these claims about the benefits of disagreement, see Mercier and Sperber (2011), and De Cruz and De Smedt (2013).



I will consider each of the two versions of the problem in turn.

### 5.1 Solving the Justification Version

According to the self-undermining justification argument, adopting conciliationism leads to a loss of justification for believing conciliationism, which is incompatible with continuing to be committed to the theory. This argument only appears sound because its fifth premise is supported by:

**Unjustified Theory:** If a subject *S* is not justified in believing (or would not be justified in believing) a theory, then *S* is rationally required to give up (not be committed to) that theory.

According to the theory of endorsement, this principle is false. The attitude of commitment towards conciliationism that it is appropriate for a subject to have is endorsement, not belief. It can be rational to endorse a theory even when one would be unjustified in believing it. Rationally endorsing a theory requires less evidential support than justified belief. The extrinsic epistemic reasons that bear on whether to endorse a theory can provide strong support for endorsing it, while providing no support for believing it. Indeed, one of the main benefits of endorsement is that it vindicates commitments to theories which have inadequate evidential support to justify belief in them. Committed advocacy of a theory is beneficial to inquiry, so we have good reason to positively evaluate researchers who engage in this practice, and to actively encourage such a practice.

Once we recognize the value of endorsement, it is clear that we should not be committed to the *Unjustified Theory Principle*. Without this commitment, the self-undermining justification argument is unsound. The reason why the argument seemed compelling to many people is that *Unjustified Theory* is intuitive and plausible, as long as we think of belief as the only kind of commitment one could take toward a theory. However, once we recognize endorsement, we can see the principle is false.

Recall *Connie the Conciliator* from section 3. Connie's case can help illustrate how this solution works. Connie endorses conciliationism. Now, after engaging with the literature and discovering actual disagreements with some peers in the epistemology community, CV requires her to lower her confidence in CV. Now she has a confidence too low to be compatible with justified belief. But that's fine! She doesn't believe her theory in the first place; she endorses it. And endorsement is compatible with a relatively low confidence. Connie has good extrinsic epistemic reasons to maintain her endorsement. Doing so creates a better distribution of labor in epistemology, promotes beneficial disagreement, and helps avoid premature consensus. So Connie is well-justified in endorsing the theory, even though she would be unjustified in believing it. Her lack of justified belief is no longer reason to give up the theory. Thus, Connie's justification for being a conciliationist is not undermined. The self-undermining justification problem is defused by endorsement.

One might worry that there remains a problem for conciliationism here. Perhaps we could give a version of the self-undermining justification argument in terms of belief, swapping in “believes” for each instance of “adopting,” and “does not believe” for “giving up”. The resulting argument is sound: clearly, if one is unjustified in believing a theory, one should not believe the theory. Is this not the deeper problem: that one cannot believe that conciliationism is true? The short answer, I think, is no. The conciliationist should not view this as a problem, but rather an expected consequence that fits the spirit of the theory. CV is no worse off in this regard than any other theory from a cutting-edge research field. What would be a problem is if one could never, in principle, rationally believe conciliationism. For the reasons mentioned above (section 3), the self-undermining objection does not even allege that this is the case. It is only a problem for believing the theory here and now, in light of present disagreement. But since endorsement is the proper attitude for such a view anyway, it is no special problem for conciliationism that only endorsement (and not belief) is warranted toward CV.

## 5.2 Solving the Inconsistency version

The self-undermining inconsistency objection admits of a similar solution. Recall that, according to the inconsistency version of the problem, conciliationism offers inconsistent advice in cases of disagreement, since it recommends increased confidence in a competitor that offers conflicting advice. Since CV recommends greater confidence in SV, it thereby also recommends steadfastness as a rule. And it is impossible to both stay fast and conciliate.

As I argued above, however, the inconsistency problem also requires a commitment to an additional epistemic principle, over and above commitment to CV. It requires commitment to:

**Most Credence:** If a subject has the most credence in a theory, then the subject should follow any rules prescribed by that theory.

The *Most Credence Principle* is an initially attractive principle for governing our decision-making under uncertainty. However, the theory of endorsement gives us strong, independent reason to deny it. The principle is incompatible with the kind of resilient commitment that endorsement involves, which provides significant benefits to inquiry. It would lead to subjects giving up theories too quickly, and without adequate exploration and defense. This is because, in cutting-edge fields of inquiry, new evidence is constantly being accumulated. Subjects who obey the *Most Credence Principle* will fail to be resilient in their endorsements, and this will be harmful for inquiry.

Without appeal to some enkratic principle like the *Most Credence Principle*, the self-undermining inconsistency argument is unsound. All the versions of such principles which have been appealed to in setting up self-undermining objections are incompatible with healthy inquiry (because they are incompatible with endorsement). Thus, the argument for inconsistency is unsound. This

solution to self-undermining inconsistency can be illustrated by returning to Connie. Here, we will have to consider two versions of the case.

Again, Connie endorses conciliationism. When she is faced with disagreement about disagreement, CV recommends lowering her confidence in CV, and raising her confidence in SV. However, given that *Most Credence* is false, CV's prescriptions do not constrain her response: her high credence in CV does not mean she is rationally obligated to follow CV's prescription. She thus has two reasonable options: she can remain steadfast or lower her confidence in CV.<sup>32</sup> Neither option will result in self-undermining. Let's consider each in turn, as a proof-by-cases argument.

On the first option, Connie remains steadfast in her confidence in CV. Again, this option is permissible because *Most Credence* is false. Moreover, this option seems reasonable: CV is a contentious thesis in cutting-edge field characterized by pervasive disagreement. Someone motivated by conciliationist intuitions might well be careful about following general rules prescribed by such a contentious principle.<sup>33</sup> Despite ignoring its prescription here, Connie continues to endorse the theory, as she has good extrinsic epistemic reason to do so. Thus, her confidence permissibly remains the same, and this is perfectly compatible with her continued endorsement of CV. Therefore, on this option, her commitment to CV is not undermined. She is not obligated to follow CV's prescriptions, and she does not do so. Nor does she have reason to follow the dictates of SV, so her rational commitments do not offer any conflicting advice.

On the second option, Connie lowers her confidence the way CV recommends, lowering her credence in CV (and raising it in SV). She is not obligated to do so by her commitment to CV, but the falsity of *Most Credence* does not suggest that lowering confidence in this case is impermissible. Furthermore, conciliation seems intuitively appropriate to her in this case. Now she has higher confidence in SV than CV, but this is perfectly compatible with her continued endorsement of CV. After all, she has a variety of extrinsic epistemic reasons in favor of maintaining her endorsement of conciliationism. Since *most credence* is false, however, she is not obligated by her new high credence in SV to follow rules prescribed by SV. So, conciliationism does not give conflicting advice in her case. Connie can consistently endorse CV.

On either option, Connie can proceed in a rationally permissible way. There is no rational inconsistency in the advice her rational commitments offer her. This is because *Most Credence* is false, as becomes clear once we recognize the importance of endorsement. Thus, endorsement solves this version of the self-undermining problem, too.<sup>34</sup>

<sup>32</sup>Strictly speaking, there is a third option: she could increase her confidence in CV. But this would obviously not be self-undermining.

<sup>33</sup>Pittard (2015) makes a related point, though we disagree about the precise implications.

<sup>34</sup>Note that one cannot save the self-undermining inconsistency argument by moving to talk of belief rather than adopting/giving up. This is because premise 4 of the argument is still necessary, and it is still false (as is the generalization, the *Most Credence Principle*). Moreover, adding a premise such as "follow SV if you believe it," won't help, as that still requires justification by a principle such as "if you have most credence in P, you should believe P." And I think that prin-

### 5.3 The Endorsement Principle

The endorsement solution to self-undermining works by providing independent reason to deny the epistemic principles used in the objection. One might worry that there is nonetheless some undiscovered enkratic principle that would support the self-undermining argument. However, I think there is good reason to doubt that this is the case. It's hard to see what principle would both support self-undermining and allow for adequate resilience. Moreover, for those who are inclined to think there must be *some* theory-following enkratic principle which governs inquiry, the theory of endorsement itself motivates such a principle, which I will call the *Endorsement Principle*. However, it should be stressed that the solution to self-undermining provided by endorsement does not depend on this principle.

The endorsement framework prescribes a significant commitment to base one's research program on the theory endorsed. This involves giving the theory a full exploration, which includes investigating the implications of following the rules contained in the theory. This motivates a different kind of enkratic principle for theory-following:

**Endorsement Principle:** In a research domain  $d$ , if a subject endorses a theory, then (while engaged in research in  $d$ ) the subject should follow any rules prescribed by that theory (as long as the consequences of this rule following are mostly limited to  $d$ ).

This principle is implied by endorsement's role in shaping a subject's research program. It is part of what it means to pursue and explore a theory, by investigating all of its consequences. Some of those consequences will concern how research is conducted. Thus, the Endorsement Principle helps make endorsement beneficial to inquiry. This provides significant independent motivation for this principle, as it is derived from aspects of the endorsement theory which were not designed merely to solve the self-undermining problem. The solution to self-undermining, augmented by appeal to the endorsement principle, can be illustrated by returning to Connie's case.

Connie endorses conciliationism. When she encounters disagreement about disagreement, CV recommends lowering her confidence in CV, and raising her confidence in SV. According to the endorsement principle, she should follow this recommendation and lower her credence in CV (and raise it in SV). But having higher credence in SV than CV is compatible with her continued endorsement of CV. After all, she has a variety of extrinsic epistemic reasons in favor of maintaining her endorsement of conciliationism. Since she continues to endorse CV and not SV, the endorsement principle does not recommend that she follow rules prescribed by SV. Therefore, conciliationism does not give conflicting advice in her case. According to CV and the theory of endorsement,

---

ciple is also false, or at least highly controversial. However, even if such an argument could be given, the same considerations apply to it as were raised for the belief version of self-undermining justification in the last section.

Connie should continue endorsing CV, and following its prescriptions to conciliate in the face of disagreement. So, Connie consistently endorses CV.

The final parenthetical caveat to the endorsement principle is crucial. The committed advocacy of endorsement is only appropriate when it is limited specifically to research domains. This is why I suggested above that endorsement is a fragmented attitude. One should not base practical action or public policy advice on a theory one merely endorses. Doing so would risk bad consequences, as there are good reasons to endorse theories that are probably false. It would be a bad idea to engineer bridges and buildings on such shaky foundations. The endorsement principle, though, suggests actions be taken on the basis of endorsements. This only makes sense if the (significant) consequences of these actions are limited to the research domain. Of course, I have left some vagueness in this characterization, as what counts as being mostly limited to a research domain will depend on the circumstances.

The most important upshot of this caveat for the present discussion, however, is that application of the *Endorsement Principle* to CV would require that changes to a subject's credences are limited to the research domain. This means that credences must also be fragmented attitudes. In this case, that means that if a subject lowers her credence in some theory or statement because of her endorsement of CV, this lowered credence must itself be limited to the research domain in the epistemology of disagreement. Otherwise, these changes could affect the subject's actions outside of doing epistemology research, which would be irresponsible. This suggestion, that credences are fragmented and should be compartmentalized to a particular domain, is not novel. This fragmentation reflects the limitations of human information processing. It helps explain phenomena like recall failures, systematic inconsistency, and mathematical discovery.<sup>35</sup>

In sum, the theory of endorsement suggests an enkratic principle, the *Endorsement Principle*, which is a plausible alternative to the principles we have considered and discarded. It does not lead to self-undermining. This principle is a plausible addition to the theory of endorsement, as long as one is comfortable endorsing fragmentation of credences. However, the solution to the self-undermining problem does not require the endorsement principle.

The picture of disagreement that emerges from this discussion looks like this: Connie and her peers disagree in *credence*, and it is this disagreement that CV prescribes a response to. Notice that it is not the disagreement in *endorsement* that requires Connie to lower her confidence. This means that Connie must be able to get information about what her peers' credences are. However, this is already required by any theory which embraces CV. Recognizing that a practice of endorsement is in place may actually help researchers like Connie to disambiguate what information is being provided by bare assertions of theories within research domains, and motivate them to inquire more carefully

<sup>35</sup>For more on fragmentation and its solutions to various problems, see Egan (2008); Elga and Rayo (2019); Lewis (1982); Rayo (2013); Stalnaker (1984).

about her peers' level of confidence. While it is true that researchers will need to be careful to acquire information about their peers' credences, and not just what they endorse, this is not a novel requirement introduced by the theory of endorsement.

Before discussing two significant objections to the endorsement solution to the self-undermining problem, I want to briefly mention one of its strengths. There is no space here to compare this solution in detail to other attempted solutions. However, the advantages mentioned here go some way toward making the case that this is the best solution on offer. The advantage of the solution is the fact that it is embedded in an independently motivated research program. The theory of endorsement offers a number of benefits beyond its solution to self-undermining. We need it to properly explain and to vindicate the behavior of many highly effective researchers. There are a number of additional reasons, beyond disagreement, to think that belief cannot play the role of the attitude of pursuit during inquiry. Moreover, as was argued above, the practice of endorsement confers a number of benefits to a field of inquiry. In addition, the theory helps solve a number of separate problems in social epistemology and the general philosophy of science. It promotes an appropriate distribution of cognitive labor, it helps avoid premature consensus, and it contributes to our understanding of theory change and theory pursuit. This is also the reason why it is worth introducing the endorsement framework in this paper: it offers independent motivation for denying the background assumptions that are required for both versions of the self-undermining objection.

Thus, the endorsement solution to self-undermining is embedded in an independently motivated research program. It also solves the problem in a way that respects the intuitions that motivate conciliationism, without carving out a special, ad hoc exception for the theory of conciliationism itself. These are significant advantages of the solution.

## 6 Response to Objections

Before concluding, there are two significant objections that I want to briefly address. The first objection concerns the need to appeal to the endorsement framework in order to solve the self-undermining objection. The thought is that we can accept that we should not believe conciliationism, on the basis of self-undermining objections. However, it might be rational for someone to continue to advocate for it, and to live by it, in virtue of purely pragmatic reasons. On this picture, what I call extrinsic epistemic reasons concerning the good of inquiry can be construed instead in a pragmatic manner. They are not epistemic reasons to endorse some theory, they are practical reasons to undertake certain actions, in particular, to advocate and defend the theory.<sup>36</sup>

I think it is a benefit of the analysis of self-undermining offered above that it also suggests this pragmatic response. This response depends on using prag-

---

<sup>36</sup>Thanks to an anonymous reviewer for raising this objection.

matic reasons to motivate denying the identified enkratic principles. Moreover, the pragmatic reasons in question concern what will lead to successful inquiry, so they are the same token reasons identified as extrinsic reasons by my account. So, a pragmatist can benefit from the above analysis.

Still, I want to avoid this kind of appeal to pragmatic reasons. I think it is a significant cost to the conciliationist to give up on a genuinely epistemic defense of the theory. The steadfast has no similar need to appeal to pragmatic reasons to save SV from itself. Unless one is already committed to pragmatism about belief and epistemic normativity, this seems like an embarrassment for the conciliationist. The endorsement solution offers a genuinely epistemic response. It suggests that reasons having to do with the good of inquiry are genuine epistemic reasons (of a special type), as discussed above in section 4.2. It accepts that there can be epistemic reasons for action.<sup>37</sup> There is not space for a full defense of this account of epistemic normativity here. However, if this account is right, it gives us good reason to prefer the endorsement solution to the pragmatic alternative.

The second objection I will call the *ordinary cases objection*. This objection alleges that, even if my analysis in the preceding section is correct, there is a part of the initial self-undermining objection that remains. Specifically, this worry concerns whether the conciliationist can adopt their theory, while remaining a conciliator in everyday life. The basic idea is that the endorsement solution only applies to the domain of inquiry. It protects the rationality of committing to conciliationism during inquiry, but it fails to justify a commitment to conciliationism that is relevant outside the epistemology of disagreement. That is, it fails to justify a commitment to conciliationism in ordinary cases outside of inquiry, in a way that leads to counter-intuitive consequences. The objection can be illustrated with another case:

**Mental Math Connie:** Connie proceeds as before, endorsing conciliationism, even after lowering her credence in it when faced with disagreement. Now Connie is at a restaurant with her friend, and they need to split the check. Connie calculates half the check to be \$43, while her friend (who has the same evidence and the same level of competence at elementary arithmetic) calculates it to be \$45.

What should Connie do, faced with this instance of disagreement? She endorses conciliationism, but she has a low credence in it, where “low credence” means the same as it does above: less than is required for justified belief. How should she now respond to this objection? A proponent of the objection suggests that, intuitively, she should conciliate. More importantly, they might say, any solution to the self-undermining problem should protect Connie’s ability

<sup>37</sup>This last idea is far from unprecedented. Sosa (2015) sees epistemology as a kind of performance evaluation. Hookway (2006) suggests activity is the primary locus of epistemic evaluation. Singer and Aronowitz (in press) suggest that there can be epistemic reasons for even quite ordinary actions.

to rationally conciliate in these circumstances. But the endorsement solution to self-undermining doesn't provide justification for adhering to conciliationism here. Connie's endorsement of conciliationism rationally governs her behavior within the domain of inquiry: while she is arguing in the philosophy conference, seminar, or in a journal article. By the lights of the theory of endorsement, it does not (and should not) govern her behavior outside of this domain. Instead, it is her credences which should determine her behavior in ordinary situations of uncertainty, and according to my solution, her credence in conciliationism may be quite low; perhaps even less than .5.

Thus, the ordinary cases objection alleges, there is something missing from the endorsement solution. It allows the researcher to be a committed advocate of conciliationism, but it does not protect conciliationism from being rendered un-followable in domains outside of inquiry. The endorsement solution doesn't solve the whole problem.

The response to the ordinary cases objection involves distinguishing between a) being committed to conciliationism and b) conciliating in particular circumstances. The basic idea is this: one can have good reason to conciliate in particular circumstances, even if one does not believe (or is not committed) to the theory of conciliationism. This can be true even if the theory of conciliationism is false.

Conciliationism is a theory. (More carefully, as noted above, it is a set of theories committed to CV, but we can set aside that complication for the moment). It is a general theory about how one should respond to disagreement. This theory is meant to explain our intuitions in a variety of cases, and to give us prescriptions in a wide range of more troublesome cases about which our intuitions are unclear or in conflict. This makes it much like theories in other normative domains, such as (normative) ethics. That is to say, the methodology here is to seek reflective equilibrium. Notice that our judgments, or intuitions, about particular cases such as *Mental Math* are being treated as *evidence* to be explained (at least defeasibly). That is why the theory seeks to explain them. We are generally, and I think appropriately, more confident that our intuitions are correct in these cases than we are that any general theory about them is correct. We would need very good theoretical reason to overturn these judgments.

Even more importantly, the particular, situation-specific reasons present in cases like *Mental Math* do not necessarily depend on the truth of the general theory of conciliationism.<sup>38</sup> CV may be false, while Connie's reasons to lower her confidence in her answer remain. Specifically, she has the same reasons to lower her confidence that were cited above in explaining the intuitions that motivate conciliationism in the first place. That is, the disagreement shows that either Connie or her friend has made a mistake, and Connie has no reason

<sup>38</sup>Certain extreme forms of steadfastism might not count the apparent situation-specific reasons as reasons. The *Right Reasons* view is a radical externalist version of SV that has this result (Kelly 2005; Titelbaum 2015). But even proponents of such views recognize the intuitive pull of these apparent reasons, and seek to explain them away. Moreover, someone like Connie would need to be convinced that right reasons is the correct theory, and come to believe it, before it could guide her actions.



to think it is more likely to be her than her friend. Thus, Connie need not believe conciliationism in order to justifiably believe she has reason to conciliate in this case.

The relationship between the situation-specific reasons and the general theory of conciliationism can be illustrated by analogy with theories in normative ethics. Suppose I have low confidence in, and do not believe, utilitarianism. The fact that some action in a particular circumstance will increase the aggregate happiness is still often a good reason to do that action. Consider a case in which I am deciding whether to buy apples or oranges for someone, and I find out that they would be happier if I buy apples. That it would increase overall happiness is a reason to buy the apples. I am much more confident this is true than I am of any particular theory of normative ethics. Moreover, I don't need to believe utilitarianism to think this is true, and it is likely to be true even if the theory is false.<sup>39</sup>

The point of all this is that Connie can have good reason to conciliate, even if she does not believe conciliationism, and even if it is not her commitment to the theory which motivates her change of confidence. Her conciliating can be motivated by situation-specific reasons. She need not be motivated by a commitment to conciliationism in order to conciliate.

This discussion again raises a question about which enkratic principles are plausible. I have argued that endorsement provides the appropriate normative framework for inquiry. This includes determining the correct enkratic principles in research domains. Outside of research, however, other principles must be provided. Here is a plausible candidate:

**Justified Theory Principle** Outside of research domains, if a subject has a justified belief in a theory, then the subject should follow any rules prescribed by that theory.

This principle is essentially the contrapositive of *Unjustified Theory* above, except that it only constrains subjects when they are outside of research domains. It is thus compatible with the endorsement principle.

Notice, however, that this principle does not constrain Connie, as long as her confidence in conciliationism has not been forced so low that it is compatible with a justified categorical belief in steadfastism. In Connie's mental math case, she has a categorical belief in neither conciliationism nor steadfastism, so the *Justified Theory Principle* does not apply. Connie has no relevant justified belief in a theory, and so the antecedent of the principle is not satisfied.

Furthermore, as noted, Connie has significant first-order, situation specific reasons to lower her confidence in light of disagreement in this case. Since she is not confident enough in any general principle about disagreement to be constrained to follow its dictates, and since she has significant reasons to conciliate that are particular to this case, she should conciliate. Thus, the ordinary cases objection is blocked: even those with low confidence in CV may permissibly

<sup>39</sup>Thanks to an anonymous referee for raising the worry in terms of utilitarianism.

conciliate in specific cases where that is, intuitively, what they should do. Disagreement about CV only threatens one's ability to be practically committed to the view as a general principle.<sup>40</sup>

One might still worry, on this response, that one cannot be practically committed to conciliationism. But this is what the conciliationist should expect: after all, we conciliationists think that disagreement should lower our confidence in theories, and there is too much disagreement about conciliationism for it to be reasonable for us to apply it, across the board, as a rule for decisions outside of research domains. Hopefully, someday we will reach consensus in the field, and then we will be able to so apply it. But conciliationism is no worse off in this regard than any other theory in a cutting-edge research field. In the meantime, the self-undermining objection offers no reason not to conciliate in the cases where it is clearly intuitive to do so, such as Connie's.

## 7 Conclusion

Conciliationism is a highly plausible theory in the epistemology of disagreement. However, its proponents face an objection meant to force us to give it up: the self-undermining problem. The theory of endorsement, however, offers a way out of this problem. Endorsement is the appropriate attitude to take toward the theory of conciliationism during inquiry. The theory of endorsement rationalizes the committed advocacy of researchers toward their views, even after disagreement requires that they lose confidence in them. It also justifies a distinct enkratic principle, and justifies the denial of some that have been taken for granted in the literature. This explains how one can stably and consistently maintain a rational commitment to conciliationism. Moreover, the success of this solution provides further evidence for endorsement, and its accompanying norms of inclusive epistemic rationality.

**Acknowledgments** For helpful feedback on this paper, I would like to thank Austin Andrews, D Black, Laura Callahan, Nevin Climenhaga, Marilie Coetsee, Andy Egan, Adam Elga, Megan Feeney, Melissa Fusco, James Gullledge, Alan Hájek, Branden Fitelson, Ben Henke, Brian McLaughlin, Casey O'Callaghan, Emily Prychitko, Pamela Robinson, Susanna Schellenberg, Ernest Sosa, James Willoughby, and audiences at the Australian National University and the 2019 Pacific Division meeting of the APA.

## References

Barnett, Z. (2019). Philosophy without belief. *Mind*, 128(509), 109–138. doi: 10.1093/mind/fzw076

---

<sup>40</sup>Compare this to Lasonen-Aarnio's appeal to particularism about higher-order defeat in (2013; 2014). She thinks that no enkratic principles are plausible, and so beliefs about theories or rules need have no effect at all on actions in these cases. In contrast, I do think some such rules are required, but not just any will do.

- Berker, S. (2013). Epistemic teleology and the separateness of propositions. *Philosophical Review*, 122(3), 337–393. doi: 10.1215/00318108-2087645
- Bogardus, T. (2009). A Vindication of the Equal Weight View. *Episteme*, 6(3), 324–335. doi: 10.3366/e1742360009000744
- Broome, J. (2013). *Rationality Through Reasoning*. Wiley-Blackwell.
- Christensen, D. (2007). Epistemology of Disagreement: The Good News. *Philosophical Review*, 116(2), 187–217.
- Christensen, D. (2009, September). Disagreement as Evidence: The Epistemology of Controversy. *Philosophy Compass*, 4(5), 756–767. doi: 10.1111/j.1747-9991.2009.00237.x
- Christensen, D. (2010a, July). Higher-Order Evidence1. *Philosophy and Phenomenological Research*, 81(1), 185–215. doi: 10.1111/j.1933-1592.2010.00366.x
- Christensen, D. (2010b). Rational Reflection. *Philosophical Perspectives*, 24(1), 121–140.
- Cohen, L. J. (1989). Belief and acceptance. *Mind*, 98(391), 367–389. doi: 10.1093/mind/XCVIII.391.367
- Decker, J. (2014). Conciliation and Self-Incrimination. *Erkenntnis*, 79(5), 1099–1134.
- De Cruz, H., & De Smedt, J. (2013). The value of epistemic disagreement in scientific practice: The case of homo Floresiensis. *Studies in History and Philosophy of Science Part A*, 44(2), 169–177. doi: 10.1016/j.shpsa.2013.02.002
- Egan, A. (2008). Seeing and believing: Perception, belief formation and the divided mind. *Philosophical Studies*, 140(1), 47–63.
- Elga, A. (2007, September). Reflection and Disagreement. *Noûs*, 41(3), 478–502. doi: 10.1111/j.1468-0068.2007.00656.x
- Elga, A. (2010). How to disagree about how to disagree. In R. Feldman & T. Warfield (Eds.), *Disagreement*. Oxford University Press.
- Elga, A., & Rayo, A. (2019). Fragmentation and information access. *Manuscript*.
- Elgin, C. Z. (2010). Persistent disagreement. In R. Feldman & T. A. Warfield (Eds.), *Disagreement*. Oxford: Oxford University Press.
- Firth, R. (1981). Epistemic merit, intrinsic and instrumental. *Proceedings and Addresses of the American Philosophical Association*, 55(1), 5–23. doi: 10.2307/3131397
- Fleisher, W. (2018). Rational endorsement. *Philosophical Studies*, 175(10), 2649–2675. doi: 10.1007/s11098-017-0976-4
- Fleisher, W. (2019). Endorsement and Assertion. *Noûs*, forthcoming. doi: 10.1111/nous.12315
- Frances, B., & Matheson, J. (2018). Disagreement. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Spring 2018 ed.). Metaphysics Research Lab, Stanford University.
- Friedman, J. (2017). Why Suspend Judging? *Noûs*, 51(2), 302–326.
- Gardiner, G. (2014, March). The Commutativity of Evidence: A Problem for Conciliatory Views of Peer Disagreement. *Episteme*, 11(1), 83–95. doi: 10.1017/epi.2013.42
- Goldberg, S. (2013). Defending philosophy in the face of systematic disagreement. In D. Machuca (Ed.), *Disagreement and skepticism* (pp. 277–94). New York: Routledge.
- Goldman, A. I. (1987, October). Foundations of social epistemics. *Synthese*, 73(1), 109–144. doi: 10.1007/BF00485444
- Greaves, H. (2013). Epistemic decision theory. *Mind*, 122(488), 915–952. doi: 10.1093/mind/fzt090
- Hookway, C. (2006). Epistemology and Inquiry: The Primacy of Practice. In S. Hetherington (Ed.), *Epistemology Futures* (pp. 95–110). Oxford University Press.

- Horowitz, S. (2014). Epistemic akrasia. *Noûs*, 48(4), 718–744.
- Jehle, D., & Fitelson, B. (2009). What is the “equal weight view”? *Episteme*, 6(03), 280–293.
- Jenkins, C. (2007). Entitlement and rationality. *Synthese*, 157(1), 25–45. doi: 10.1007/s11229-006-0012-2
- Kelly, T. (2005). The Epistemic Significance of Disagreement. In J. Hawthorne & T. Gendler (Eds.), *Oxford Studies in Epistemology, Volume 1* (pp. 167–196). Oxford University Press.
- Kelly, T. (2010). Peer Disagreement and Higher Order Evidence. In A. I. Goldman & D. Whitcomb (Eds.), *Social Epistemology: Essential Readings* (pp. 183–217). Oxford University Press.
- King, N. L. (2012, September). Disagreement: What’s the Problem? or A Good Peer is Hard to Find. *Philosophy and Phenomenological Research*, 85(2), 249–272. doi: 10.1111/j.1933-1592.2010.00441.x
- Kitcher, P. (1990). The division of cognitive labor. *The Journal of Philosophy*, 87(1), 5–22. doi: 10.2307/2026796
- Lackey, J. (2008). A Justificationist View of Disagreement’s Epistemic Significance. In A. M. A. Haddock & D. Pritchard (Eds.), *Proceedings of the Xxii World Congress of Philosophy* (pp. 145–154). Oxford University Press.
- Lasonen-Aarnio, M. (2010). Unreasonable Knowledge. *Philosophical Perspectives*, 24(1), 1–21.
- Lasonen-Aarnio, M. (2013, December). Disagreement and Evidential Attenuation. *Noûs*, 47(4), 767–794. doi: 10.1111/nous.12050
- Lasonen-Aarnio, M. (2014, March). Higher-Order Evidence and the Limits of Defeat. *Philosophy and Phenomenological Research*, 88(2), 314–345. doi: 10.1111/phpr.12090
- Laudan, L. (1978). *Progress and its problems: Towards a theory of scientific growth*. Berkeley: Univ of California Press.
- Laudan, R. (1987). The rationality of entertainment and pursuit. In *Rational changes in science* (pp. 203–220). Dordrecht: Springer.
- Levi, I. (1980). *The enterprise of knowledge: An essay on knowledge, credal probability, and chance*. Cambridge, MA: MIT Press.
- Levinstein, B. A. (2015, March). With all due respect: The macro-epistemology of disagreement. *Philosophers Imprint*, 15(13), 1–20.
- Lewis, D. (1971). Immodest Inductive Methods. *Philosophy of Science*, 38(1), 54–63.
- Lewis, D. (1982). Logic for Equivocators. *Noûs*, 16(3), 431–441.
- Longino, H. E. (1990). *Science as social knowledge: Values and objectivity in scientific inquiry*. Princeton: Princeton University Press.
- Lougheed, K., & Simpson, R. M. (2017). Indirect epistemic reasons and religious belief. *Religious Studies*, 53(2), 151–169. doi: 10.1017/S0034412516000202
- Maher, P. (1993). *Betting on theories*. Cambridge: Cambridge University Press.
- Matheson, J. (2015a, April). Are Conciliatory Views of Disagreement Self-Defeating? *Social Epistemology*, 29(2), 145–159. doi: 10.1080/02691728.2014.907833
- Matheson, J. (2015b, April). Are Conciliatory Views of Disagreement Self-Defeating? *Social Epistemology*, 29(2), 145–159. doi: 10.1080/02691728.2014.907833
- Matheson, J. (2015c). Disagreement and Epistemic Peers. *Oxford Handbooks Online*. doi: 10.1093/oxfordhb/9780199935314.013.13
- McKaughan, D. (2007). *Toward a richer vocabulary for epistemic attitudes* (Unpublished doctoral dissertation). University of Notre Dame.
- McKaughan, D. (2008). From Ugly Duckling to Swan: C. S. Peirce, Abduction, and the

- Pursuit of Scientific Theories. *Transactions of the Charles S. Peirce Society*, 44(3), 446–468.
- McMullin, E. (1976). The Fertility of Theory and the Unit for Appraisal in Science. In R. S. Cohen, P. K. Feyerabend, & M. Wartofsky (Eds.), *Essays in Memory of Imre Lakatos* (pp. 395–432). Reidel.
- Mercier, H., & Sperber, D. (2011, April). Argumentation: Its adaptiveness and efficacy. *Behavioral and Brain Sciences*, 34(2), 94–111. doi: 10.1017/S0140525X10003031
- Mulligan, T. (2015). Disagreement, Peerhood, and Three Paradoxes of Conciliationism. *Synthese*, 192(1), 67–78. doi: 10.1007/s11229-014-0553-8
- Nickles, T. (1981). What is a Problem That We May Solve It? *Synthese*, 47(1), 85–118.
- Pettigrew, R. (2016). *Accuracy and the Laws of Credence*. Oxford University Press Uk.
- Pittard, J. (2015, January). Resolute Conciliationism. *The Philosophical Quarterly*, 65(260), 442–463. doi: 10.1093/pq/pqv022
- Plantinga, A. (1999). Pluralism: A Defense of Religious Exclusivism. In K. Meeker & P. Quinn (Eds.), *The Philosophical Challenge of Religious Diversity* (pp. 172–192). New York: Oxford University Press.
- Psillos, S. (1999). *Scientific realism: How science tracks truth* (1edition ed.). London ; New York: Routledge.
- Rayo, A. (2013). *The Construction of Logical Space*. Oxford University Press.
- Rinard, S. (2015). No Exception for Belief. *Philosophy and Phenomenological Research*, 91(2).
- Rosenkranz, S., & Schulz, M. (2015). Peer Disagreement: A Call for the Revision of Prior Probabilities. *Dialectica*, 69(4), 551–586.
- Schoenfield, M. (2018). An Accuracy Based Approach to Higher Order Evidence. *Philosophy and Phenomenological Research*, 96(3), 690–715. doi: 10.1111/phpr.12329
- Šešelja, D., Kosolovsky, L., & Straßer, C. (2012). The Rationality of Scientific Reasoning in the Context of Pursuit: Drawing Appropriate Distinctions. *Philosophica*, 86, 51–82.
- Šešelja, D., & Straßer, C. (2013). Kuhn and the Question of Pursuit Worthiness. *Topoi*, 32(1), 9–19.
- Šešelja, D., & Straßer, C. (2014). Epistemic justification in the context of pursuit: A coherentist approach. *Synthese*, 191(13), 3111–3141. doi: 10.1007/s11229-014-0476-4
- Shah, N., & Velleman, J. D. (2005). Doxastic Deliberation. *The Philosophical Review*, 114(4), 497–534.
- Singer, D. J., & Aronowitz, S. (in press). What Epistemic Reasons Are For: Against the Belief-Sandwich Distinction. In B. Dunaway & D. Plunkett (Eds.), *Meaning, Decision, and Norms: Themes from the Work of Allan Gibbard*.
- Solomon, M. (1994). Social empiricism. *Noûs*, 28(3), 325–343. doi: 10.2307/2216062
- Sorensen, R. (1988). *Blindspots*. Oxford: Clarendon Press.
- Sosa, E. (2010). The Epistemology of Disagreement. In A. Haddock, A. Millar, & D. Pritchard (Eds.), *Social Epistemology*. Oxford University Press.
- Sosa, E. (2015). *Judgment and agency*. Oxford: Oxford University Press.
- Staffel, J. (2019). Credences and Suspended Judgments as Transitional Attitudes. *Philosophical Issues*, 29(1), 281–294. doi: 10.1111/phis.12154
- Stalnaker, R. (1984). *Inquiry*. Cambridge: Cambridge University Press.
- Steel, D. (2010). Epistemic values and the argument from inductive risk. *Philosophy of Science*, 77(1), 14–34. doi: 10.1086/650206
- Strevens, M. (2003, February). The role of the priority rule in science. *The Journal of Philosophy*, 100(2), 55–79. doi: 10.5840/jphil2003100224

- Titelbaum, M. (2015). Rationality's fixed point (or: In defense of right reason). *Oxford studies in epistemology*, 5, 253–94.
- Van Fraassen, B. C. (1980). *The scientific image*. Oxford; New York: Clarendon Press ; Oxford University Press.
- Weatherston, B. (2013). Disagreements, Philosophical and Otherwise. In J. Lackey & D. Christensen (Eds.), *The Epistemology of Disagreement: New Essays* (p. 54). Oxford University Press.
- Weintraub, R. (2013). Can Steadfast Peer Disagreement Be Rational? *Philosophical Quarterly*, 63(253), 740–759. doi: 10.1111/1467-9213.12065
- Whitt, L. A. (1990). Theory pursuit: Between discovery and acceptance. *PSA: Proceedings of the biennial meeting of the philosophy of science association*, 1, 467–483.
- Whitt, L. A. (1992). Indices of theory promise. *Philosophy of Science*, 59(4), 612–634. doi: 10.1086/289698
- Wiland, E. (2018). Peer Disagreement: Special Cases. *Logos and Episteme*, 9(2), 221–226. doi: 10.5840/logos-episteme20189217
- Worsnip, A. (2014). Disagreement About Disagreement? What Disagreement About Disagreement? *Philosophers' Imprint*, 14.