

Peer Disagreement and Group Inquiry

Jeff Dunn
jeffreydunn@depauw.edu

1 Introduction

Suppose that you and I regard each other as epistemic peers about some topic. Suppose that we find that we disagree about proposition P . Perhaps I believe it and you disbelieve it. More subtly, perhaps I assign P a credence of 0.9 while you assign it a credence of 0.1. According to *Conciliationism*, once we learn of our disagreement, we should move our opinions in P in the direction of each other.¹

In a recent paper defending this view, David Christensen (2007b) writes:

In general, I think it's true that if those who work in poor epistemic conditions were more epistemically rational, there would be less disagreement, and many positions would be held with less confidence. This would, of course, have some unwelcome consequences, even beyond the frustration of acknowledging that we often cannot confidently answer the questions we study. I think there would even be a possible epistemic downside to this sort of result. It's quite plausible that knowledge is best advanced by people exploring, and attempting to defend, a variety of answers to a given question. Perhaps, human psychology makes this easier to do when investigators actually have a lot of confidence in the hypotheses they're trying to defend. Certain sorts of inquiry might well work best when a variety of investigators have irrationally high levels of confidence in a variety of pet hypotheses. So there may well be important epistemic benefits to

¹Numerous authors have defended some version of this view, including Feldman (2006), Elga (2007, 2010), Christensen (2007b, 2011), Bogardus (2009), Matheson (2009), Kornblith (2010), and van Wietmarschen (2013). Other authors are committed to a view that says at least some of the time we should move our opinions in this way, including the Justificationist View due to Jennifer Lackey (2010) and the Total Evidence View due to Thomas Kelly (2010). My argument is mainly directed toward the stronger version of Conciliationism, but does affect these other views.

certain patterns of irrational belief. But I would argue that the patterns of belief are no more epistemically rational for all that.
(p. 215)

In this paper, I seek to do two things. First, I will present psychological evidence and a formal model that together support what Christensen finds plausible: that in certain cases of peer disagreement, inquiry is best served by participants in the disagreement *not* moderating their opinions—at least not for some time. Second, I will argue that this is more problematic for Conciliationism than Christensen here maintains: if there are known epistemic benefits to not being conciliatory this raises problems for Conciliationism.

2 The Initial Tension

Conciliationism says that if we regard each other as epistemic peers on a topic and then learn of a disagreement about a proposition within that topic, each of us should move our opinion on the disputed proposition in the direction of the other. There are several points of clarification needed.

First, what is it to regard another as an epistemic peer on a topic? The rough idea is that we are epistemic peers with respect to a topic when we share a common body of total evidence about the topic, we have equal cognitive ability with respect to the topic, and these abilities are not temporarily impaired. So, we regard each other as peers when we believe that these conditions are satisfied. Adam Elga (2007) gives a different definition: I count you as my epistemic peer on a proposition P iff conditional on finding out that we disagree about P , I think it is equally likely that each of us is mistaken. Note that Elga's definition is different than the rough definition in terms of shared evidence and cognitive ability since I could think that we're equally likely to be mistaken conditional on a disagreement about P even if I know that we do not share the same evidence or cognitive ability. I'll adopt something like Elga's definition in this paper. This is because the motivating idea behind Conciliationism in cases where peers believe they share the same evidence and cognitive ability extends to cases where they know they don't share the same evidence or cognitive ability, but still regard it as equally likely that they each are mistaken. For instance, suppose that you and I find that we disagree with each other on a certain proposition within our shared area of expertise. Suppose that we can confidently rule out that we have exactly the same evidence and exactly the same reasoning abilities. Nevertheless, we are uncertain whose set of evidence or whose reasoning abilities are superior. As King (2012) urges, those committed to

Conciliationism should say that learning of our dispute in such a case still gives you and I a reason to withhold our beliefs.

One final comment about epistemic peer-hood is necessary. Because some of the early papers on peer disagreement focussed on protracted philosophical disagreements (of the sort between, for instance, David Lewis and Peter van Inwagen on free will (van Inwagen, 1996)), it is tempting to think that epistemic peers are those who are thoroughly acquainted with each others' views and evidence, and who have engaged in extended debate with each other. Though that's one way to be an epistemic peer, it's not the only way. In this paper, I'll focus on epistemic peers who have engaged in no debate with each other about the disputed proposition. For all this, however, they can still be epistemic peers; they can still regard the other person an expert in the area under dispute and think that given the disagreement, it is roughly equally likely that either one of them is mistaken.²

The second point of clarification concerns what it means to move one's opinion toward one's peer. A perfectly general answer to this question is difficult, but there are particularly clear cases. If we are working with all-or-nothing belief, then if I believe P and you believe not- P , we move our opinions toward each other by withholding belief in P and not- P . If we are working with graded belief, then if my credence in P is 0.9 and your credence in P is 0.1, then we move toward each other by moving toward 0.5. How *much* need we modify our credences? This is a complicated question.³ But in many cases, Conciliationism will recommend that we move to very similar credences near 0.5. This is certainly what Christensen has in mind in the quote in the previous section.

Finally, Conciliationism says that we *should* move our opinions in the direction of each other. But what is the nature of this normative claim? For some authors it seems to denote an all-things-considered epistemic obligation.⁴ For others, the 'should' in Conciliationism only denotes that the moderated opinions are *justified*, *rational*, or *best supported by our evidence*.⁵

²For further discussion of the issue of epistemic peer-hood see Christensen (2011) and Comesaña (2012).

³For a nice discussion of some of the complexities, see Fitelson & Jehle (2009).

⁴For instance, Adam Elga (2010) writes: "According to such views, finding out that a respected advisor disagrees with one should move one at least a little in the direction of the the advisor's view." (p. 176) Ballantyne & Coffman (2012) give a similar definition: "In a revealed peer disagreement over P , each thinker should give at least some weight to her peer's attitude." (p. 657).

⁵Thomas Kelly (2013), for instance, defines the view as follows: "... a view counts as Conciliationist if and only if it entails that suspending judgment is a necessary condition for being reasonable in a canonical case of peer disagreement." (p. 36) David Christensen

Of course, these views may not come apart. One might hold a view according to which the opinions one is epistemically obligated to adopt just are the opinions that are justified, or rational, or best supported by the evidence. The important point is that Conciliationism may differ depending on the answer one gives to this normative question. In section 6 I address this question directly.

If that's Conciliationism, then how is it in tension with evidence about group inquiry? Here's the basic idea. There is evidence that groups of inquirers answer questions more accurately when group members defend different viewpoints than when all the members defend the same view. Further, the evidence suggests that this increase in accuracy occurs only if group members actually *believe* the views they defend. If our overriding epistemic goal is to believe the truth, then this seems to give us pretty strong reason to not be conciliatory. Consider an example. Scientists from Lab A defend theory *A* and those from Lab B defend competing theory *B*. The psychological data suggest that they are more likely to find the true theory if Lab A members go on believing and defending *A*, and Lab B members go on believing and defending *B*. With respect to the *A/B* controversy, the lab members should believe whatever will make it likely they find the true theory. So, the lab members should go on believing and defending their respective theories. Conciliationism, however, says that this is incorrect: with respect to the *A/B* controversy that lab members should moderate their opinions toward each other.

There are, of course, quick ways to resolve this tension. One can, for instance, take the line that Christensen does in the quote that opens this paper. One can maintain that in some sense it serves the lab members' epistemic ends to defend their respective views, but that for all that, they are not being epistemically rational in so-doing. I'll consider this kind of response in detail, but will first present empirical evidence and a formal model that together support the claim that conciliatory responses to disagreement are not always the most truth-conducive.

3 Empirical Evidence

In this section I'll argue that there is empirical evidence that supports:

(P) Groups reach more accurate conclusions (1) when there is genuine dis-

(2007b, 2011) also makes clear that he sees Conciliationism as concerning epistemic rationality. Along these lines, Jonathan Matheson (2009) defines Conciliationism as concerning justification.

sent and debate than when there is not, *and* (2) when the dissenting parties really hold dissenting beliefs about the debated proposition(s).

Start with (1). First, there is evidence that groups can outperform any individual member in reasoning tasks. Consider a study by Moshman & Geil (1998). In the study, the participants were divided into 3 experimental conditions: individual control, interactive condition, individual/interactive condition. In the individual control, the participants were asked to solve the Wason Selection Task on their own.⁶ In the interactive condition, participants were asked to solve the task in groups with 5-6 members. In the individual/interactive condition, participants were first asked to solve the task alone, and then (without having the correct answer revealed to them) solve the task in a group. The results are striking. In the individual control condition, consistent with other studies on the Wason Selection Task, the success rate was 9.4%. In the interactive condition, the success rate jumped to 70%. Finally, in the individual/interactive condition, when these individuals worked in *groups*, the success rate was 80%.⁷

Groups can perform better than any individual member. There is, however, and important caveat: for groups to do better than individuals, group members must be debating and arguing with each other in a genuine way. This fits well with the phenomenon of group polarization. Sunstein (2002), for instance, presents evidence that when individuals that share a view on an issue discuss that issue, this leads the individuals to *increase* their confidence in the shared view. In this context, group deliberation will not do any more than polarize the antecedently held view. So, *debate*—not mere *discussion* with like-minded individuals—is important to attain the benefits of group deliberation.

This point has been emphasized in a series of recent papers by Mercier & Sperber (Mercier & Sperber, 2011; Mercier, 2012, 2011). They argue for the *argumentative theory of reasoning*, which maintains that reasoning ability developed in our evolutionary past to facilitate arguing. This hypothesis is controversial. However, a large swath of the evidence that the argumentative theory of reasoning can explain is direct evidence in favor of (1). Mercier (2011) writes:

⁶See Wason (1966), Johnson-Laird & Wason (1970), Wason & Shapiro (1971). For those unfamiliar with the task, the participant is presented with four cards. It is stipulated that each card has a number on one side and a letter on the other. The participant can see the following faces of the card: 4, K, 7, E. The participant is then asked which cards must be flipped over to confirm that the following rule is being followed: *If a card has an even number on one side, then it has a vowel on the other.*

⁷For similar studies, see Michaelsen *et al.* (1989) and Stasson & Bradshaw (1995).

... reasoning should be at its best in argumentive contexts. Such contexts naturally arise when a group willing to work together disagrees—this is the kind of context for which reasoning, it has been suggested, evolved. Reasoning should be activated more easily and should be more efficient in such contexts, much in the same way as color vision is at its best in broad daylight. There is now a wealth of evidence supporting this hypothesis. (p. 183)

Similarly, Mercier & Sperber (2011) write: “...many experiments have shown that *debates* are essential to any improvement of performance in group settings.” (p. 63, my emphasis).

To get a feel for the studies that support the idea that genuine debate is needed to reap the advantages of group problem solving, consider Schulz-Hardt *et al.* (2006). The study involved groups attempting to solve *hidden profile* problems. These are problems where the correct solution requires some set of information, but no one group member possesses this set of information. For instance, the problem might be to select the best apartment. In a hidden profile problem, the full set of information clearly picks out one apartment as best, but each group member only has partial information. Groups discuss the problem and then decide on a group answer. Schulz-Hardt *et al.* manipulated the information available to each group member to test the effect of diversity of opinion on the eventual solution the groups put forward. Groups where all group members had full information reached the correct decision 100% of the time. Groups with homogenous preferences for options before discussion reached the correct decision 7% of the time. Groups where group members had diverse preferences before discussion—but no member who initially preferred the correct option—reached the correct decision 26% of the time. Finally, groups where group members had diverse preferences before discussion—and at least one member who initially preferred the correct option—reached the correct decision 62% of the time. Overall, groups with heterogenous views reached the correct decision 43% of the time.⁸ It seems, then, that groups with dissent can reach accurate answers more often than either the individual group members or groups that do not have dissent.⁹

⁸See also Brodbeck *et al.* (2002), Perret-Clermont *et al.* (2004), and Kuhn *et al.* (1997).

⁹Gigone & Hastie (1997) express some skepticism about the robustness of the phenomenon that groups are more accurate than individuals. They note that certain conditions often must be met for groups to outperform the members individually. In particular, they refer to earlier work (Hastie, 1986, pp. 151-2) that identifies three characteristics that produce high levels of group performance: (1) the problem has a “eureka solution”, a solution that may not be obvious initially but is demonstrable once discovered, (2) in-

So, why do groups that engage in dissent tend to do better? Mercier (2012) suggests one intriguing idea. There is evidence that we are very good at looking for arguments that confirm a view we have already settled on. This is related to the phenomenon known as *confirmation bias*.¹⁰ There is also evidence that we are good at looking for falsifying arguments or data, but only for views with which we disagree.¹¹ In light of this, a heterogeneous group will uncover a more balanced set of evidence than a homogenous group.

While this empirical data is interesting, it is natural to object at this point that these studies aren't relevant to Conciliationism. Why? Well, within the peer disagreement literature epistemic peers are usually defined to be those who have the same evidence relative to a certain question. But in the studies I've just cited, the disagreement is contrived by providing group members with different evidence. There are two lines of response to this worry. First, as already mentioned (section 2), although epistemic peers are usually defined to be those who share the same evidence, the motivations for Conciliationism will yield advice to be conciliatory even in cases where the disagreeing parties know they do not share the same evidence but are unsure about whose evidence is superior. The second response is that other studies (to be discussed shortly) suggest that we weigh evidence differently depending on the beliefs we have. One plausible explanation for why heterogeneous groups see increases in accuracy is that in heterogeneous groups the evidence gets a "fair hearing" because it is debated by those who are biased in complementary ways. Thus, though disagreement is contrived in experimental situations by giving group members different evidence, there is reason to think the advantages of heterogeneity would still function in cases where group members share the same evidence.

The data so far canvassed suggest that groups of inquirers that engage in genuine debate about possible answers to a question are ultimately more accurate in answering a question than groups of inquirers that do not engage in such debate. This does not yet show that group members must *believe* the propositions for which they argue. Perhaps groups could get all the advantages of genuine debate even though no group member *believes* the views he or she is defending. This would allow all the group members

dividual judgment accuracy is perturbed by unsystematic errors, and (3) group members possess different evidence. All this is entirely consistent with what I maintain here: that in certain situations, groups are more accurate than individuals.

¹⁰See Nickerson (1996) for general information on confirmation bias. See also Evans (1996).

¹¹See, for instance, Dawson *et al.* (2002).

to be conciliatory in their beliefs and yet still accrue the advantages of heterogenous group inquiry. I now argue that there is evidence in favor of (2) and so against this idea.

Some evidence for (2) is provided by the well-known phenomenon of *belief bias*.¹² In experimental situations, participants are asked to evaluate arguments that agree with their beliefs (no-conflict) and arguments that disagree with their beliefs (conflict). For example, a no-conflict case would require a participant to diagnose an invalid argument with an absurd conclusion as invalid or a valid argument with a true conclusion as valid. In a conflict case, a participant must diagnose a valid argument with an absurd conclusion as valid or an invalid argument with a true conclusion as invalid. Experimental subjects perform much better in no-conflict cases than in conflict cases. This, in turn, suggests that we are better at reasoning when our reasoning agrees with our beliefs than when it does not. Thus, merely playing the devil's advocate is *harder* than defending a view one sincerely believes.

Further research into belief bias provides more evidence for (2). De Neys & Franssens (2009) show that those who do better in the conflict cases are those who are able to inhibit their beliefs on the matter, at least for a time. Interestingly, they also show that when people fail to do well on conflict cases it is not because they do not attempt to inhibit their belief, but rather because they attempt to inhibit their belief unsuccessfully. This is evidence that we're not good at inhibiting our beliefs in a way that yields argumentative advantages, even if we sincerely try.

More direct evidence for (2) is provided by Schulz-Hardt *et al.* (2002). They investigated whether various groups that consisted of genuinely disagreeing members perform similarly to groups where dissent is contrived using various devil's advocacy techniques. They found that groups that had genuine disagreement were less biased in seeking only confirmatory information.¹³ More evidence for this comes from Greitemeyer *et al.* (2006). In their study they set out to show that, contrary to the results just described, contrived dissent *can* yield advantages for group reasoning. They instructed various group members to defend different points of view, even if that point of view was not one the group member really believed. Though group members did play the appropriate roles, and as a result a more balanced menu of evidence was discussed, the group answers did not improve. In contrast,

¹²There is considerable evidence for this phenomenon. For the seminal study, see Evans *et al.* (1983).

¹³Strauss *et al.* (2011) summarizes this study as well as other related ones.

in the genuinely heterogenous groups, group answers did improve markedly.

What could explain why real heterogeneity rather than contrived heterogeneity is required to improve decision making? Greitemeyer *et al.* (2006) suggest an interesting hypothesis. We tend to weigh evidence for a proposition differently depending on whether we believe the proposition or not. For example, if we believe that government stimulus is required in response to a recession, then even if we are exposed to balanced evidence, we will give greater weight to the evidence in favor of the efficacy of government stimulus and discount the dissenting evidence. Thus, in groups where dissent is contrived, more evidence may come out, but it is not taken seriously if the members have homogenous beliefs.

Putting all this evidence together provides good reason to think that group members who are investigating some question will end up with more accurate beliefs about the answer to that question if they initially maintain their divergent beliefs and vigorously defend them. Of course, this is not true in *every* instance of group inquiry. If the group members are to reap the benefits of this group discussion, they must eventually converge on one answer. And the members must engage in genuine debate with each other. And, of course, group members must adopt the group decision at the end of inquiry if they are to individually reap this benefit. But when these conditions are met, the goal of accuracy is served best in group inquiry when group members maintain their divergent beliefs and defend them.

4 Formal Model

The empirical claim, (P), is fairly well established. But it doesn't have any normative consequences on its own, being a purely descriptive claim about what happens in group inquiry. Thus, (P) alone doesn't cause any problems for Conciliationism. For it to cause trouble we need some sort of principle that bridges accuracy to epistemic normativity.

Here's a promising principle to fill that gap. Suppose that S has two epistemic options, A and B, with respect to some question Q. The principle says:

(E): If the expected accuracy (given S's evidence) of option A with respect to question Q is greater than the expected accuracy (given S's evidence) of option B with respect to Q, then it is not the case that S (epistemically) should take option B rather than A (relative to Q).

First note that (E) is formulated in terms of *expected* accuracy rather than accuracy. Why? This is because we want a principle that will bridge

accuracy to epistemic normativity where the normativity is of the same sort that features in Conciliationism. But the ‘should’ in Conciliationism is definitely not one that tracks accuracy directly; otherwise, it would make no sense to say that disagreeing peers should withhold belief. In so-doing Conciliationism would be saying that one should sometimes give up an accurate belief. One might have doubts whether the appeal to expected accuracy does any better in making contact with Conciliationism—something I consider in section 6—but at least it has a shot.

How does (E) apply to the situation of disagreeing peers? Let’s regiment our case. First, we assume that our group is one that will engage in debate about the issue and where each group member is disposed to adopt the group consensus at the end of inquiry. Initially our group members come into a group meeting with a belief about the correct answer to Q. At t_1 they notice that other group members disagree with them about which answer to Q is correct. At this time, each group member has the option of remaining *steadfast* in her belief about the correct answer to Q or withholding belief about the correct answer to Q and thereby being *conciliatory*. Then, the group engages in debate, which is eventually concluded at t_2 with the group reaching some consensus.

There are several ways we can think about evaluating the expected accuracy of either remaining steadfast or being conciliatory at t_1 . One simple approach looks at the expected accuracy only at the end of inquiry, at t_2 . Another approach looks at the expected accuracy at the end of inquiry (t_2), but also the expected accuracy when the decision is made at t_1 . In fact, there is a whole family of approaches here where we weight the expected accuracies at t_1 and t_2 in different ways. Giving full weight to t_2 corresponds to only focusing on the end of inquiry. Giving full weight to t_1 corresponds to not being concerned with future accuracy at all. And between these extremes are many other options. Two natural options include the approach where each time is given equal weight, and the approach where t_2 is given more weight than t_1 . This latter approach is plausible if the time between t_1 and t_2 is not great, but the answer adopted at t_2 will be held for a long time. The formal model below will show that the option of being steadfast often comes out with a higher expected accuracy than being conciliatory so long as we give each time equal weight or give greater weight to t_2 . Thus, principle (E) says that agents in such groups should not be conciliatory.

To show this, we need to make things slightly more rigorous. First, we need to say something precise about the epistemic options that agents have.

I will focus on options that consist solely of belief states adopted.¹⁴ So, the options—or epistemic acts—that are available to S at some time will include only belief states that S can come to occupy at that time. What kinds of belief states should we allow? There are two salient approaches. According to the first approach, beliefs are all-or-nothing: you either believe, disbelieve, or withhold belief in a proposition. According to the second approach, beliefs are graded: you can adopt any strength of belief in a proposition between full belief and full disbelief. I will model belief as graded, although in analyzing the situation of peer disagreement I will simplify things considerably by allowing graded beliefs to take only three values: one value that will correspond to belief, one that will correspond to disbelief, and one that will correspond to withholding belief.

Suppose, then, that the question under investigation has a finite number m of mutually exclusive and exhaustive possible answers in $\Omega = \{1, \dots, m\}$. Let a belief state with respect to this question, \mathbf{c} , be a probability vector $\{c_1, \dots, c_m\}$ such that $c_1, \dots, c_m \geq 0, c_1 + \dots + c_m = 1$. Thus, c_1 is the degree of belief—or *credence*—assigned to answer 1, c_2 the credence assigned to answer 2, etc. We want to evaluate these credences for how accurate they are. The appropriate formal tool for this is a *scoring rule*, a function $S(\mathbf{c}, i)$ that tells us how close to the truth \mathbf{c} is in world i .¹⁵ Scoring rules thus give scores to credences based on only two things: the level of credence assigned to the answer and what the answer is.

There are many scoring rules that can be used. I'll focus on the popular Brier score, which is given by:

$$BS(\mathbf{c}, i) = \sum_{j=1}^m (\delta_{ij} - c_j)^2,$$

where $\delta_{ij} = 1$ if $i = j$ and $\delta_{ij} = 0$ otherwise. Note that the Brier Score has a minimum value of zero and a maximum value that increases with increasing m .¹⁶ According to the Brier Score, lower numbers are better. Thus, we can see the Brier Score as measuring *inaccuracy*: less is better.

¹⁴Hilary Greaves (2013) makes this a condition of an act being an *epistemic act*.

¹⁵More precisely, if \mathcal{P}_m is the set of all probability vectors of length m , then a scoring rule is a function $S(\mathbf{c}, i) : \mathcal{P}_m \rightarrow \mathbb{R}, i = 1, \dots, m$. Other conditions need to be placed on scoring rules to yield ones that can serve the purpose of measuring *accuracy*. For more on these conditions, see Joyce (2009). Instead of go through these here, I'll adopt the popular Brier score.

¹⁶If it is important to compare scores between questions that have different numbers of possible answers we can normalize the Brier Score by taking $BS(\mathbf{c}, i)/m$ where m is the number of possible answers to the question.

Once we have our scoring rule in place, we can work out the expected inaccuracy of adopting some particular belief state, \mathbf{c} . A natural way to do this is to see how well \mathbf{c} expects itself to do. This yields:

$$ES(\mathbf{c}) = \sum_i c_i \times S(\mathbf{c}, i),$$

for $c_i \in \mathbf{c}$. For several reasons, however, this is inadequate for our purposes. First, given the structure of our scenario, we need to be able to evaluate the adoption of belief state \mathbf{c} at some time for its effects on accuracy at a later time. The proposal above, however, does not easily allow this since the c_i that serve as weights for the expectation only range over possible answers to the question. Second, we'd like to be able to evaluate the expected value of adopting belief state, \mathbf{c} , using weights that might differ from the agent's own perspective of the situation. In particular, we'd like to be able to use a probability function, p , that tracks the evidential support that various hypotheses have in light of the agent's current evidence even if the agent's belief state doesn't match this.¹⁷ The ability to have this mismatch is important in this context since we want to be open to the idea that it is reasonable for an agent to adopt a belief state that is not best supported by her evidence.

We can address both of these shortcomings by working with the following expected score:

$$ExS(A, t_n) = \sum_i \sum_{\mathbf{c}_{t_n}} p(i \wedge \mathbf{c}_{t_n} | A) \times S(\mathbf{c}_{t_n}, i).$$

A is some epistemic act, such as the option to remain steadfast or be conciliatory. These acts will have consequences for the agent's current and perhaps future belief states, but they need not dictate with certainty one particular belief state at those times. t_n is the time at which the consequences are being evaluated (note that this can be, but need not be, the same time at which option A is taken). \mathbf{c}_{t_n} are the credences adopted by the agent at t_n . Finally, p is a probability function that meets the description in the previous paragraph: it tracks the evidential support for hypotheses based on the agent's evidence *now*.¹⁸ So, $ExS(A, t_n)$ gives us the expected inaccuracy of option A where we are only concerned with the consequences at t_n . So,

¹⁷Ralph Wedgwood (ms) argues for something similar to this, albeit in a different context.

¹⁸It's worth mentioning that ExS is an epistemic analogue of *evidential* expected utility rather than *causal* expected utility. While I prefer causal decision theory over evidential decision theory, this is a distinction that won't matter in this context. In particular, we

principle (E) says that S should not choose option B over option A when the sum of $ExS(A, t_n)$ over all times t_n relevant to the decision is less than the sum of $ExS(B, t_n)$.¹⁹ In the next section I'll apply this principle to the case of peer disagreement.

5 Against Conciliationism

We want to evaluate two different epistemic acts:

S : the act of remaining steadfast at t_1

C : the option of being conciliatory at t_1 .

To keep things simple, we'll assume that our question has two possible answers $\Omega = \{P, \neg P\}$ and that there are only three belief states that one can adopt: $\{c_P = 1, c_{\neg P} = 0\}$, $\{c_P = 0, c_{\neg P} = 1\}$, and $\{c_P = 0.5, c_{\neg P} = 0.5\}$. These correspond, respectively, to believing P , believing $\neg P$, and withholding belief with respect to P . Denote these with ' c^B ', ' c^D ', ' c^W '.²⁰ We'll assume that taking act C results in withholding belief at t_1 and that taking act S results in either believing P or believing $\neg P$ at t_1 .

In the following three subsections I show the following things:

- If we evaluate the expected inaccuracy only at t_1 , expected inaccuracy is minimized by withholding belief (at t_1) upon learning of a peer disagreement.
- If we evaluate the expected inaccuracy only at t_2 , given the empirical data about homogenous vs. heterogenous groups, expected inaccuracy is minimized by remaining steadfast (at t_1) upon learning of a peer disagreement.
- For individuals who belong to certain groups, the (unweighted) sum of the expected inaccuracy at both times (t_1 and t_2) is minimized by remaining steadfast at t_1 . (Note that this means that any weighting of the sum that gives greater weight to t_2 than to t_1 will also give the verdict that remaining steadfast at t_1 minimizes expected inaccuracy.)

won't be dealing with cases where an epistemic act is evidentially but not causally relevant to outcomes. So, for the sake of simplicity, I'll work with ExS . For more on this, see Greaves (2013).

¹⁹The sum of the preferred act should be less than the alternatives since our score measures *inaccuracy*.

²⁰In general, if a question has m possible answers, the option of withholding can be modeled as an assignment of value $1/m$ to each possible answer.

These results, together with (E), imply that individuals in certain groups should not be conciliatory in cases of peer disagreement.

For those interested in the formal details of these results, sections 5.1–5.3 provide them. However, if one wants to skip the formal details, one can skip to section 6 without loss. Section 6 is where I consider the philosophical question about whether these results pose a genuine problem for Conciliationism.

5.1 Time of Evaluation: t_1

We want to calculate the expected inaccuracy of choosing S or C at t_1 with the consequences evaluated at t_1 . That is, we'd like to work out:

$$ExS(S, t_1) = \sum_i \sum_{\mathbf{c}_{t_1}} p(i \wedge \mathbf{c}_{t_1} | S) \times S(\mathbf{c}_{t_1}, i).$$

$$ExS(C, t_1) = \sum_i \sum_{\mathbf{c}_{t_1}} p(i \wedge \mathbf{c}_{t_1} | C) \times S(\mathbf{c}_{t_1}, i).$$

Since choosing act C will result at t_1 in belief state $\mathbf{c}_{t_1}^W$, it follows that $p(\mathbf{c}_{t_1}^W | C) = 1$ and so $p(P \wedge \mathbf{c}_{t_1}^W | C) = p(P | C)$ and $p(P \wedge \mathbf{c}_{t_1}^B | C) = p(P \wedge \mathbf{c}_{t_1}^D | C) = 0$ (and similarly for ' P ' replaced with ' $\neg P$ '). It is plausible that my choosing to be conciliatory does not affect the probability that P is true. Thus, $p(P | C) = p(P)$. Thus, $ExS(C, t_1)$ can be simplified to:

$$ExS(C, t_1) = \sum_i p(i) \times S(\mathbf{c}_{t_1}^W, i),$$

where i can take value either P or $\neg P$.

A similar line of argument simplifies $ExS(S, t_1)$ to either of the following, depending on whether the agent in question either initially believed or disbelieved P :

$$ExS(S, t_1) = \sum_i p(i) \times S(\mathbf{c}_{t_1}^B, i)$$

$$ExS(S, t_1) = \sum_i p(i) \times S(\mathbf{c}_{t_1}^D, i),$$

where i can take value either P or $\neg P$. If our scoring rule is symmetric—in that it assigns the same penalty to $c(P) = n$ when P is true as it assigns to $c(P) = 1 - n$ when P is false—then these two expected scores will be equal.

Note that these formulas are specific instances of a general class. The class contains those problems where (i) the epistemic acts under consideration are precise credal values that can be assigned to elements in Ω , and

(ii) where the fact that I've adopted some credal assignment isn't relevant to which element in Ω obtains. In such situations, it is in general true that:

$$ExS(\mathbf{c}_{t_1}, t_1) = \sum_i p(i) \times S(\mathbf{c}_{t_1}, i)$$

So long as S is a proper scoring rule²¹, as the Brier Score is, this will be minimized for $c_i = p(i)$. What this means is that if we focus on specific beliefs adopted at a time, evaluated for their epistemic consequences at that time, we get that one should always believe in accordance with what the evidence supports (which is encoded by p). This is one way in which the approach adopted here seems to be on the right track: many think that epistemic rationality consists in believing in accordance with what your evidence supports now. Here we have a decision-theoretic rationale for this.²²

To see how this plays out in the case of group inquiry we must settle one more thing: the value of $p(P)$ at t_1 . It is plausible and also charitable to Conciliationism to hold that it is at or near 0.5. This is because at the time of the decision we have an equal number of equally qualified individuals, half of whom believe P and the other half who believe $\neg P$. Those who defend Conciliationism have given various arguments to support the claim that in such situations, the evidence supports neither P nor $\neg P$. If we normalize the Brier scores, then under these conditions we have that $ExBS(C, t_1) = 0.25$ and $ExBS(S, t_1) = 0.5$.

Summing Up: If we look only at the consequences at t_1 , an agent minimizes expected inaccuracy by being conciliatory at t_1 .

5.2 Time of Evaluation: t_2

We want to calculate the expected value of choosing at t_1 either C or S but with the consequences now evaluated at t_2 , after the group has debated the issue. That is, we'd like to work out:

$$ExS(S, t_2) = \sum_i \sum_{\mathbf{c}_{t_2}} p(i \wedge \mathbf{c}_{t_2} | S) \times S(\mathbf{c}_{t_2}, i).$$

²¹A proper scoring rule is a scoring rule for credences that has the following property: the credence function that has the best expected score from the perspective of any coherent credence function, c , is c itself. For more on this see Seidenfeld (1985).

²²It is important to note, however, that this holds only in the case where the fact that I've adopted some credal assignment isn't relevant to which element in Ω obtains. If that fails, then we can get some scenarios where one is advised to believe against the evidence. For more on this, see Greaves (2013), Caie (2013), and Carr (unpublished).

$$ExS(C, t_2) = \sum_i \sum_{\mathbf{c}_{t_2}} p(i \wedge \mathbf{c}_{t_2} | C) \times S(\mathbf{c}_{t_2}, i).$$

What is important here are the values we assign to the probability function, p . Recall that we are allowing \mathbf{c}_{t_2} to take only three forms: $\mathbf{c}_{t_2}^B$, $\mathbf{c}_{t_2}^W$, and $\mathbf{c}_{t_2}^D$. ($P \wedge \mathbf{c}_{t_2}^B$) and ($\neg P \wedge \mathbf{c}_{t_2}^D$) can thus be thought of as adopting a true belief at t_2 ; ($P \wedge \mathbf{c}_{t_2}^W$) and ($\neg P \wedge \mathbf{c}_{t_2}^W$) can be thought of as withholding belief at t_2 ; and ($P \wedge \mathbf{c}_{t_2}^D$) and ($\neg P \wedge \mathbf{c}_{t_2}^B$) can be thought of as holding a false belief at t_2 . The question about the probability function is whether any of these outcomes are more likely conditional on the agent either being conciliatory or steadfast at t_1 , which is what S and C denote.

The psychological data suggest that when a group is attempting to reach consensus about some proposition, they will be considerably less accurate in the consensus reached if all the members have the same belief about the proposition than if they have different beliefs. If the whole group is conciliatory, then they will have the same belief concerning P , whereas if they are all steadfast, then they will have different beliefs. Accordingly, the evidence suggests that conciliatory groups tend to be less accurate than steadfast groups. However, this doesn't yet settle *our* question. For the probabilities we want to know are the not probabilities of accuracy conditional on the *whole group* choosing S or C . They are probabilities of accuracy conditional on *one* group member choosing S or C , and the empirical literature doesn't take a stand on what these probabilities might be.

Nevertheless, it is relevant to the question. To see why, we need to take note of a general probabilistic fact. The fact is this: If Y screens off X with respect to Z in the sense that $p(Z|X \wedge Y) = p(Z|Y)$ and $p(Z|X \wedge \neg Y) = p(Z|\neg Y)$, then $p(Z|X) = p(Z|Y)p(Y|X) + p(Z|\neg Y)p(\neg Y|X)$.²³ This allows us to execute a kind of probabilistic hypothetical syllogism. Why is this helpful? Well, we want to know $p(i \wedge \mathbf{c}_{t_2}|S)$ and likewise with S replaced by C . Let GS stand for the proposition that the *group* is steadfast and GC stand for the proposition that the *group* is conciliatory. Assuming that I adopt the group consensus at the end of inquiry, the empirical literature gives us some idea about the value of $p(i \wedge \mathbf{c}_{t_2}|GS)$ and $p(i \wedge \mathbf{c}_{t_2}|GC)$. In particular, we know that a steadfast group is more likely to be accurate than a conciliatory group. If we had some idea about $p(GS|S)$ and if GS screens off S with respect to whether the belief at t_2 is true, false, or withheld, then we could use this to work out $p(i \wedge \mathbf{c}_{t_2}|S)$.

²³Shogenji (2003).

First, note that facts about whether the group is conciliatory or steadfast do screen off facts about what I do with respect to whether the consensus of the group is accurate. Thus, we can apply the probabilistic fact yielding:

$$p(i \wedge \mathbf{c}_{t_2}|S) = p(i \wedge \mathbf{c}_{t_2}|GS)p(GS|S) + p(i \wedge \mathbf{c}_{t_2}|\neg GS)p(\neg GS|S),$$

and similarly with ‘ S ’ replaced with ‘ C ’ and ‘ GS ’ replaced with ‘ GC ’.

Thus, all we must do now is to determine the values of $p(GS|S)$ and $p(\neg GS|S)$ as well as $p(GC|C)$ and $p(\neg GC|C)$. In general, *my* particular decision might make little difference to the group, especially if it is a large group. However, in some special circumstances, we might be able to get some insight into these probabilities. Focus on a scenario that is popular in the peer disagreement literature: the scenario where there are only two group members, you and me, and we come into the dispute with different beliefs: I believe P and you believe $\neg P$. If we both remain steadfast, then we’ll get the benefit of real debate; if we’re both conciliatory, then we won’t. But what if we focus just on my decision? To say something about this, we need to say something about when a group counts as having heterogenous beliefs about P in the way that yields the benefits of group inquiry. On one way of going, we could say that whenever you and I assign significantly different credences to P , we hold different beliefs that will yield the benefits of group inquiry. If we go this way, then if I withhold belief and you remain steadfast, then we will get the benefits of group inquiry. On a different way of going, we could say that we get the benefits of group inquiry only if one of us holds very high credence in a proposition while the other one holds very low credence in a proposition. If we go this way, then if I withhold belief and you remain steadfast, then we will not get the benefits of group inquiry. The empirical literature doesn’t take a stand on which kind of different beliefs for group members yields accuracy benefits. So let’s consider each case in turn.

On the first approach, there is reason to think that *me* being steadfast raises the chance of our group being steadfast. I know that you have two options. You will either remain steadfast and hold onto $c(P) = 0$ or you will be conciliatory and adopt $c(P) = 0.5$. If you do the former, it doesn’t matter what I do; we’ll still get the benefits of group inquiry. However, if you do the latter, then if I am conciliatory too we do not reap the benefits of group inquiry. Thus, by remaining steadfast I can guarantee that we’ll get the benefits of group inquiry. On the second approach we get the same result. If I am conciliatory, then I ruin our chance to get the benefits of group inquiry. On the second approach we only get that if we hold extreme

and opposed views about P . If I'm steadfast, I cannot guarantee that we do get the benefits of group inquiry since you might choose to be conciliatory, but I leave the possibility open.

Putting this all together, let ' t_{GS} ', ' f_{GS} ', and ' w_{GS} ' denote the probability that a group reaches a true answer, a false answer, or withholds belief conditional on the group being steadfast (and similarly with ' GC ' replacing ' GS '). Assume that if a group is not steadfast, then it is conciliatory, which is plausible in two-person groups.²⁴ We want to work out the probability that the group consensus is correct/incorrect/withheld conditional on me being steadfast or being conciliatory. Denote this with ' t_S ', ' f_S ', ' w_S ', etc. Given the probabilistic fact above, we have:

$$p(\text{true belief at } t_2|S) = t_S = t_{GS} \times p(GS|S) + t_{GC} \times p(GC|S),$$

$$p(\text{true belief at } t_2|C) = t_C = t_{GC} \times p(GC|C) + t_{GS} \times p(GS|C),$$

and similarly for the other outcomes. What are the values of the conditional probabilities here? Let's adopt a simple model that corresponds with the second way of thinking about heterogeneity of group belief. In a two-person group, we each have two options: we can remain steadfast or be conciliatory. If we are both steadfast, then we have a steadfast group; otherwise we are conciliatory. Supposing that both options are equally likely for me and you, we have:

$$\begin{array}{ll} p(GS|S) = 1/2 & p(GS|C) = 0 \\ p(GC|S) = 1/2 & p(GC|C) = 1 \end{array}$$

This yields:

$$t_S = \frac{t_{GS} + t_{GC}}{2}$$

$$t_C = t_{GC}$$

and similarly for the other options.

From the empirical literature, we have that $t_{GS} > t_{GC}$. This implies that $t_S > t_C$. If we also hold that $f_{GS} < f_{GC}$, which is entailed by the first inequality if groups are not allowed to withhold belief at the end of inquiry, then we have that $f_S < f_C$. From these two suppositions, it follows that

²⁴Even with larger groups this is not to assume that groups must contain only all members who are steadfast or all who are conciliatory. Rather, it is to assume that there is some number of steadfast individuals that is necessary to spur the advantages of heterogeneous group problem solving, and that if a group does not have this necessary number, then the group is no longer steadfast and thus is a conciliatory group.

the expected inaccuracy at t_2 is always less conditional on being steadfast compared to being conciliatory.

We can relax the assumption that $f_{GS} < f_{GC}$ while still holding that $t_{GS} > t_{GC}$. This amounts to maintaining that the steadfast group is more likely than the conciliatory group to get a correct answer, though they are also more likely to form a belief that is false. This can happen if the conciliatory group is more likely than the steadfast group to simply withhold belief. In this situation, it does not always minimize inaccuracy at t_2 for me to be steadfast. If we stick with the Brier score, then remaining steadfast does best if and only if $w_{GC}/4 + f_{GC} > w_{GS}/4 + f_{GS}$. A specific situation where this holds is where $t_{GS} = 0.6$, $f_{GS} = 0.3$, $t_{GC} = 0.29$, and $f_{GC} = 0.2$.

Of course, the specific contours of the probability function, p , will depend on many things. But this suggests that for all inquirers, there is reason to think that remaining steadfast when working with one other person increases the probability of accuracy after debate.

Summing up: If we look only at the consequences at t_2 , then in a wide range of cases, an agent minimizes expected inaccuracy by being steadfast at t_1 .

5.3 Time of Evaluation: t_1 and t_2

So far we have shown that there are realistic cases where if you care about the epistemic consequences at t_1 for a belief adopted at t_1 you should believe one thing, but if you care about the epistemic consequences at t_2 for a belief adopted at t_1 you should believe another thing.

But it is tempting to think that what matters epistemically, all things considered, is one's accuracy summed up over all times. (One might have doubts: if so, hold off until section 6.) If that's what we're interested in, then what is the epistemically best thing to believe at t_1 ? Still sticking with the Brier score, we know that at t_1 , the option of remaining steadfast has expected inaccuracy of 0.5 whereas the option of being conciliatory has expected inaccuracy of 0.25. So, the question is whether the expected inaccuracy for act S taken at t_1 but evaluated at t_2 is at least 0.25 less than the expected inaccuracy of C at t_2 . It is just in case the following holds:

$$f_C + w_C/4 - 1/4 > f_S + w_S/4.$$

Put in terms of the probabilities of accuracy for *groups* this becomes:

$$f_{GC} + w_{GC}/4 - 1/4 > \frac{f_{GS} + f_{GC}}{2} + \frac{w_{GS} + w_{GC}}{8}.$$

For instance, here's one specific case where the inequality holds:

| Conciliatory | Steadfast |
|----------------|-----------------|
| $t_{GC} = 0.2$ | $t_{GS} = 0.83$ |
| $w_{GC} = 0.3$ | $w_{GS} = 0.13$ |
| $f_{GC} = 0.5$ | $f_{GS} = 0.04$ |

How plausible are these values? Initially they might seem rather implausible. In particular, the conciliatory group is anti-reliable: when they reach a consensus they are much more likely to be wrong than right, even though the question had only two possible answers, P and $\neg P$. However, there is some experimental data suggesting that homogenous groups might be anti-reliable in a way that makes these values not so implausible. For instance, in Schulz-Hardt *et al.* (2006), there were four possible answers to the questions the groups attempted to solve. The homogenous groups were correct 7% of the time, whereas the heterogenous groups were correct 43% of the time. Since the chance rate would be 25%, the homogenous group is actually anti-reliable.²⁵

There are some things about the study in Schulz-Hardt *et al.* (2006) that makes it not perfectly translatable to our model. Given the experimental design, the homogenous groups never had a group member who initially believed the correct answer. Further, the homogenous groups all prefer the same one answer to the question; they are not homogenous in withholding belief. So, there is room for doubt about whether the probability values above ever obtain.

Nevertheless, in certain groups, these numbers could be plausible. Further, we don't need numbers as dramatic as this for remaining steadfast to have a lower expected inaccuracy than being conciliatory *if* we give a greater weight to the accuracy at t_2 rather than t_1 . And independent of all that, this shows that it is possible that one should not be conciliatory at t_1 even though one's evidence supports being conciliatory. More carefully: if epistemic evaluation is concerned with the sum of the epistemic consequences at each time, then this result follows. This is interesting. It is different than any extant argument against Conciliationism and it grants the Conciliationist the claim that in a case of peer disagreement, the evidence supports withholding belief.

²⁵We could try to extend their data to a case with two possible answers by inferring that a homogenous group will perform at the chance rate times a factor of $7/25 = 0.28$, whereas a heterogenous group will perform at the chance rate times a factor of $43/25 = 1.72$. Applying this to a case with only two answers, we get that the homogenous group should be correct 14% of the time and a heterogenous group correct 86% of the time.

6 Epistemic Evaluation

I've just argued that in certain scenarios where there is a disagreement concerning P , the epistemic consequences with respect to P are worse overall if one is conciliatory rather than steadfast. This holds even granting the claim that the evidence supports being conciliatory with respect to P . But some might argue that this doesn't show anything wrong with Conciliationism because the consequentialist verdict is just *not* relevant to Conciliationism. Put another way, whatever "should" we get from such a consequentialist analysis is different than the "should" we get from Conciliationism. Here, for instance, is part of the quote from Christensen with which this paper started:

Certain sorts of inquiry might well work best when a variety of investigators have irrationally high levels of confidence in a variety of pet hypotheses. So there may well be important epistemic benefits to certain patterns of irrational belief. But I would argue that the patterns of belief are no more epistemically rational for all that. (p. 215)

Christensen here maintains that even if holding a certain belief has many important epistemic benefits, the belief may nevertheless still be epistemically irrational.²⁶

Christensen (2004) touches on the same theme. He asks us to consider an agent who is more confident than his evidence permits in his own abilities. Though irrational in one sense, this overconfidence may lead the person to more success in a variety of endeavors, including intellectual ones. It thus may contribute to his overall epistemic improvement. Christensen maintains that the overconfident beliefs would nevertheless not thereby be epistemically rational. In a footnote he clarifies his position:

I should note that nothing above is meant to deny that it might be rational in the pragmatic sense to cultivate self-aggrandizing

²⁶Richard Feldman (1988) asserts a similar view: "...given that I am in the situation I am in, and given that I am considering proposition p , what should I do—believe it, disbelieve it, or suspend judgment about it? Which of these three options is epistemically best? In thinking about these questions, one is to consider only these three options and only the end of getting at the truth about p . The particularly epistemic aspect of this is supposed to exclude from consideration other factors, such as which attitude would feel good or be comforting or be morally valuable. Also irrelevant to this judgment are the long-term epistemic consequences of adopting the belief. It is the truth about p now that matters." (p. 249)

beliefs. It might even be pragmatically rational to do this if one's practical goals were restricted to, e.g., maximizing one's confidence in true claims and minimizing one's confidence in false ones. Thus, if we discovered effective techniques for promoting overblown self-assessments. . . , these techniques might rightly be recommended by someone whose main concern was with epistemic melioration. But this point only highlights the distinction between the philosophical study of epistemic rationality and the project of general cognitive improvement. (p. 173)

This can be turned into an objection to the argument in this paper. The objection is this: all that has been shown here is that it is pragmatically rational to remain steadfast in certain disagreements when we restrict our attention to epistemic consequences. But that verdict is irrelevant to epistemic rationality and thus to Conciliationism.

In responding to this objection, consider two different ways one might evaluate S's belief that *P*. First, one could evaluate how well S's belief that *P* fits the evidence that S now has with respect to *P*. Second, one could evaluate how well S's belief that *P* leads S to be accurate with respect to *P*. If 'how well' is read in a subjective sense (say, in terms of minimizing expected inaccuracy), then these two ways of evaluating S's belief that *P* will often be in agreement. This is because it is often the case that the way to minimize expected inaccuracy with respect to *P* is to follow one's evidence with respect to *P*. In particular, consider Christensen's case of irrational overconfidence that leads to intellectual success. Believing that I am fantastically brilliant when there is no indication that I am is evaluated negatively with respect to evidence fitting since such a belief doesn't fit my evidence. But it is also evaluated negatively with respect to accuracy because believing I'm fantastically brilliant in such a situation doesn't lead me to be accurate about my level of brilliancy; rather, it leads me to be accurate about other things.

But, as this paper has shown, these two ways of evaluating beliefs will not always give the same verdict. We've seen that in certain cases, holding a steadfast belief with respect to *P* leads one to be more accurate with respect to *P* than does holding a conciliatory belief with respect to *P*. If Conciliationism were evaluating beliefs with respect to how well they garner accuracy, such cases would show Conciliationism to be mistaken. The view can be insulated from the objection raised in this paper by understanding it as solely evaluating beliefs based on how well they fit the evidence—as

Christensen puts it, evaluating them for *epistemic rationality*.²⁷

While this move does avoid one kind of problem for Conciliationism, I claim that it does come at a cost. For both accuracy-maximization and evidence-fitting are recognizably epistemic concerns. When they come apart—as they do here—it doesn’t *resolve* the issue to stipulate that Conciliationism is just about one of the concerns. Further, I’ll argue shortly that in the kinds of peer disagreements discussed in this paper, Conciliationism is either wrongly telling us that we all-things-epistemic-considered should be conciliatory or irrelevantly telling us only what our evidence supports.

Here’s the argument for this. Suppose there is a situation where you are faced with a difficult and important question *Q*, and offered two choices: you can take the accuracy pill or the evidence pill. If you take the accuracy pill, it will make it very likely that you believe truly the answer to *Q*. If you take the evidence pill, it will ensure that you have a belief in answer to *Q* that respects the evidence you now have. Now, these pills might yield the same belief. But imagine that we’re in a case where the evidence pill and the accuracy pill will give you different beliefs with respect to *Q*. In such a scenario, I think that you epistemically should take the accuracy pill. Imagine that someone takes the evidence pill instead and then defends himself: “I realize I am now likely less accurate in my answer to *Q* than I could have been, but at least I followed my evidence.” There’s no sense to be made of this. This demonstrates that when your evidence supports that believing in accordance with the evidence and believing accurately come apart, you epistemically should not believe in accordance with the evidence.²⁸

²⁷Though not focusing on Conciliationism, this seems to be Feldman’s view about reasonable belief: “The current view is that reasonable beliefs are epistemically valuable and that following one’s evidence is a perfect means to getting valuable beliefs.” (Feldman, 2000, p. 686).

²⁸In a series of papers Christensen (2007a, 2010) has pointed out how epistemic rationality sometimes requires us to become less accurate. For instance, if I know that with a 50% chance I’ve been slipped a logic-distorting drug, then even if I am supremely confident that *P* logically implies $P \vee Q$, epistemic rationality requires that I reduce my confidence that such an implication holds. Similarly, if I have in fact calculated our shares of the bill correctly as \$43, epistemic rationality requires that I reduce my confidence in this judgment when I learn that you calculated our shares as \$45. In both these cases, epistemic rationality requires us to become less accurate. Christensen defends this as a positive feature—not a problem—with his view. But notice that I do not claim that whenever epistemic rationality and considerations of accuracy pull in different directions, that one’s overriding epistemic obligation is to go with accuracy. Rather, I claim that in the specific situation where there is good evidence that one option will lead to a more accurate belief and the other to a more rational belief, then your overriding epistemic obligation is to go with accuracy.

But the kinds of peer disagreement described in this paper are of just this sort. Accordingly, in such situations, you should not be conciliatory. Conciliationism either says that you all-epistemic-things-considered should, in which case it is wrong, or it takes no stand on what you all-epistemic-things-considered should believe (instead merely telling you what your evidence supports). But in cases of peer disagreement we don't just want to know what our evidence supports, at least not once we realize it might come apart from accuracy. Christensen himself tells us that Conciliationism will address the question of what we should believe: “[It] would mandate extensive revision to our opinions on many controversial matters” (Christensen, 2011, p. 1). Perhaps the truth of Conciliationism does mandate extensive revision to our opinions in cases of peer disagreement that do not meet the conditions of group inquiry discussed here. But in this important class of peer disagreements, it seems that Conciliationism does not have the mandate Christensen suggests it does.

I don't want to suggest that there is *no* interest in the question of what our evidence supports. The arguments above attempt to convince one of the likely consequences of adopting different beliefs by appealing to some empirical work in psychology. But if we want to take advantage of this kind of work from psychology, then we need to know what the evidence in psychology supports now with respect to group problem solving. It is this evidence that we make use of in realizing that we can minimize our expected inaccuracy over time by remaining steadfast. We *do not* want to rely on belief states that don't best represent the world given our evidence now but instead promise to minimize our inaccuracy going forward. Such belief states could be highly misleading as to the way things are. What applies to this particular issue concerning the psychology of group inquiry applies more generally. To have any kind of realistic shot at maximizing anything we care about—accuracy, pleasure, money—we need a good representation of what the world is like now, given our evidence now.²⁹

So, there is certainly good reason to focus on the question of what our evidence supports. But that doesn't mean that in every situation one epis-

²⁹The story I've just told mimics the structure of an argument given by Hilary Kornblith (2002, 1993). Kornblith argues that everyone has a reason to value a cognitive system that generates true beliefs since having true beliefs about how the world works is required if one is to plan how to attain the things that one values. Beliefs produced by cognitive systems that generate beliefs solely to make one happy, in contrast, are not the kinds of beliefs one should use to plan how to attain the things that will make one happy. This explains, according to Kornblith, why we should evaluate cognitive systems by their conduciveness to truth.

temically should believe what one's evidence supports. And one of the reasons for the interest in peer disagreement is that we want to know what we epistemically should believe upon learning that a peer disagrees with us. Thus, I conclude that there is a genuine dilemma here for Conciliationism. On the one hand, it is taking a stand on what one should believe in cases of peer disagreement. It says that in all such cases, one should moderate one's opinions. If that's what it says, though, then it is challenged by the arguments above. On the other hand, Conciliationism might only be taking a stand on what the evidence supports in cases of peer disagreement. In that case it avoids any conflict with the empirical evidence. But this insulation from the empirical evidence comes at a cost: the view is now not taking a stand on one of the main issues in cases of peer disagreement: when I find that a peer disagrees with me, what should I believe? The empirical data suggests that the conciliatory answer is not always the correct one to this important epistemic question.

References

- Ballantyne, N. & Coffman, E. J. (2012). Conciliationism and uniqueness. *Australasian Journal of Philosophy*, 90(4), 657–670.
- Bogardus, T. (2009). A vindication of the equal-weight view. *Episteme*, 6(3), 324–335.
- Brodbeck, F. C., Kerschreiter, R., Mojzisch, A., Frey, D., & Schulz-Hardt, S. (2002). The dissemination of critical, unshared information in decision-making groups: The effects of pre-discussion dissent. *European Journal of Social Psychology*, 32(1), 35–56.
- Caie, M. (2013). Rational probabilistic incoherence. *Philosophical Review*, 122(4), 527–575.
- Carr, J. (unpublished). Epistemic utility theory and the aim of belief. <http://web.mit.edu/jrcarr/www/pdfs/accuracyevidence.pdf>.
- Christensen, D. (2004). *Putting Logic In Its Place*. Oxford University Press.
- Christensen, D. (2007a). Does murphy's law apply in epistemology? self-doubt and rational ideals. *Oxford Studies in Epistemology*, 2, 3–31.
- Christensen, D. (2007b). Epistemology of disagreement: the good news. *Philosophical Review*, 116, 187–217.

- Christensen, D. (2010). Higher-order evidence. *Philosophy and Phenomenological Research*, *81*, 185–215.
- Christensen, D. (2011). Disagreement, question-begging and epistemic self-criticism. *Philosophers' Imprint*, *11*(6), 1–22.
- Comesaña, J. (2012). Conciliation and peer-demotion in the epistemology of disagreement. *American Philosophical Quarterly*, *49*(3), 237–251.
- Dawson, E., Gilovich, T., & Regan, D. T. (2002). Motivated reasoning and performance on the wason selection task. *Personality and Social Psychology Bulletin*, *28*(10), 1379–1387.
- De Neys, W. & Franssens, S. (2009). Belief inhibition during thinking: Not always winning but at least taking part. *Cognition*, *113*(1), 45–61.
- Elga, A. (2007). Reflection and disagreement. *Noûs*, *41*(3), 478–502.
- Elga, A. (2010). How to disagree about how to disagree. In R. Feldman & T. Warfield (Eds.), *Disagreement*, Oxford University Press. 175–186.
- Evans, J. S. B. T. (1996). Deciding before you think: Relevance and reasoning in the selection task. *British Journal of Psychology*, *87*, 223–240.
- Evans, J. S. B. T., Barston, J. L., & Pollard, P. (1983). On the conflict between logic and belief in syllogistic reasoning. *Memory & Cognition*, *11*(3), 295–306.
- Feldman, R. (1988). Epistemic obligations. *Philosophical Perspectives*, *2*, 235–256.
- Feldman, R. (2000). The ethics of belief. *Philosophy and Phenomenological Research*, *60*(3), 667–695.
- Feldman, R. (2006). Epistemological puzzles about disagreement. In S. Hetherington (Ed.), *Epistemology Futures*, Oxford University Press. 216–326.
- Fitelson, B. & Jehle, D. (2009). What is the “equal weight view”? *Episteme*, *6*(3), 280–293.
- Gigone, D. & Hastie, R. (1997). Proper analysis of the accuracy of group judgments. *Psychological Bulletin*, *121*(1), 149–167.
- Greaves, H. (2013). Epistemic decision theory. *Mind*, *122*, 915–952.

- Greitemeyer, T., Schulz-Hardt, S., Brodbeck, F. C., & Frey, D. (2006). Information sampling and group decision making: The effects of an advocacy decision procedure and task experience. *Journal of Experimental Psychology: Applied*, *12*(1), 31–42.
- Hastie, R. (1986). Experimental evidence on group accuracy. In B. Grofman & G. Owen (Eds.), *Information Pooling and Group Decision Making: Proceedings of the Second University of California, Irvine Conference on Political Economy*, JAI Press. 129–157.
- Johnson-Laird, P. & Wason, P. (1970). A theoretical analysis of insight into a reasoning task. *Cognitive Psychology*, *1*(2), 134–148.
- Joyce, J. (2009). Accuracy and coherence: Prospects for an alethic epistemology of partial belief. In F. Huber & C. Schmidt-Petri (Eds.), *Degrees of Belief*, Springer. 263–297.
- Kelly, T. (2010). Peer disagreement and higher order evidence. In R. Feldman & T. Warfield (Eds.), *Disagreement*, Oxford University Press. 183–217.
- Kelly, T. (2013). Disagreement and the burdens of judgment. In D. Christensen & J. Lackey (Eds.), *The Epistemology of Disagreement: New Essays*, Oxford University Press. 31–53.
- King, N. L. (2012). Disagreement: What’s the problem? or a good peer is hard to find. *Philosophy and Phenomenological Research*, *85*(2), 249–272.
- Kornblith, H. (1993). Epistemic normativity. *Synthese*, *94*, 357–376.
- Kornblith, H. (2002). *Knowledge and Its Place in Nature*. Oxford University Press.
- Kornblith, H. (2010). Belief in the face of controversy. In R. Feldman & T. Warfield (Eds.), *Disagreement*, Oxford University Press. 29–53.
- Kuhn, D., Shaw, V., & Felton, M. (1997). Effects of dyadic interaction on argumentative reasoning. *Cognition and Instruction*, *15*(3), 287–315.
- Lackey, J. (2010). A justificationist view of disagreement’s epistemic significance. In A. Haddock, A. Millar, & D. Pritchard (Eds.), *Social Epistemology*, Oxford University Press. 298–325.
- Matheson, J. (2009). Conciliatory views of disagreement and higher-order evidence. *Episteme*, *6*(3), 269–279.

- Mercier, H. (2011). Reasoning serves argumentation in children. *Cognitive Development*, *26*(3), 177–191.
- Mercier, H. (2012). Looking for arguments. *Argumentation*, *26*(3), 305–324.
- Mercier, H. & Sperber, D. (2011). Why do humans reason? arguments for an argumentative theory. *Behavioral and Brain Sciences*, *34*, 57–111.
- Michaelsen, L. K., Watson, W. E., & Black, R. H. (1989). A realistic test of individual versus group consensus decision making. *Journal of Applied Psychology*, *74*(5), 834.
- Moshman, D. & Geil, M. (1998). Collaborative reasoning: Evidence for collective rationality. *Educational Psychology Papers and Publications*, 52.
- Nickerson, R. (1996). Confirmation bias: A ubiquitous phenomenon in many guises. *Review of General Psychology*, *2*(2), 175–220.
- Perret-Clermont, A.-N., Carugati, F., & Oates, J. (2004). A socio-cognitive perspective on learning and cognitive development. In J. Oates & A. Grayson (Eds.), *Cognitive and language development in children*, Wiley-Blackwell, chap. 8.
- Schulz-Hardt, S., Brodbeck, F., Mojzisch, A., Kerschreiter, R., & Frey, D. (2006). Group decision making in hidden profile situations: Dissent as a facilitator for decision quality. *Journal of Personality and Social Psychology*, *91*(6), 1080–1093.
- Schulz-Hardt, S., Jochims, M., & Frey, D. (2002). Productive conflict in group decision making: genuine and contrived dissent as strategies to counteract biased information seeking. *Organizational Behavior and Human Decision Processes*, *88*(2), 563–586.
- Seidenfeld, T. (1985). Calibration, coherence, and scoring rules. *Philosophy of Science*, *52*, 274–294.
- Shogenji, T. (2003). A condition for transitivity in probabilistic support. *The British Journal for the Philosophy of Science*, *54*(4), 613–616.
- Stasson, M. F. & Bradshaw, S. D. (1995). Explanations of individual-group performance differences what sort of "bonus" can be gained through group interaction? *Small Group Research*, *26*(2), 296–308.

- Strauss, S. G., Parker, A. M., & Bruce, J. B. (2011). The group matters: A review of processes and outcomes in intelligence analysis. *Group Dynamics: Theory, Research, and Practice*, 15(2), 128–146.
- Sunstein, C. (2002). The law of group polarization. *Journal of Political Philosophy*, 10(2), 175–195.
- van Inwagen, P. (1996). It Is Wrong, Everywhere, Always, for Anyone, to Believe Anything Upon Insufficient Evidence. In Jeff Jordan & Daniel Howard Snyder (Eds.), *Faith, Freedom and Rationality*, Savage, Maryland: Rowman and Littlefield. 137–154.
- van Wietmarschen, H. (2013). Peer disagreement, evidence, and well-groundedness. *Philosophical Review*, 122(3), 395–425.
- Wason, P. (1966). Reasoning. In B. Foss (Ed.), *New Horizons in Psychology*, Penguin Books. 135–151.
- Wason, P. & Shapiro, D. (1971). Natural and contrived experience in a reasoning problem. *Quarterly Journal of Experimental Psychology*, 23(1), 63–71.
- Wedgwood, R. (ms). The probabilistic structure of rational entitlement. <http://www-bcf.usc.edu/wedgwood/Entitlement.pdf>.