

Disagreement, Question-Begging and Epistemic Self-Criticism¹

David Christensen, Brown University

Subtleties aside, a look at the topography of the disagreement debate reveals a major fault line separating positions which are generally hospitable to maintaining one's confidence in the face of disagreement, and positions which would mandate extensive revision in our opinions on many controversial matters. Let us call positions of the first sort "Steadfast" and positions of the second sort "Conciliatory."² The fundamental theoretical difference between these two camps, it seems to me, lies in their differing attitudes toward evaluating the epistemic credentials of opinions voiced by people with whom one disagrees.

All parties hold that the proper response to learning of another's disagreement depends on one's epistemic evaluation of that person. All parties hold that one's beliefs about the other person's intelligence, intellectual diligence, acquaintance with the evidence, and freedom from bias, fatigue or intoxication are relevant to whether (and how much) that person's disagreement should occasion one's changing one's belief. The camps differ, though, on this question: In evaluating the epistemic credentials of an opinion expressed by someone who disagrees with me about a particular issue, may I make use of my own reasoning about this very issue? Clearly, to the extent that I may, it will favor Steadfastness. For the reasoning that supports my own view about the disputed matter will also support thinking that the other person has gotten this matter wrong.

To simplify the discussion, let us focus on cases where I've arrived at a certain degree of credence in P, and subsequently discover that another person has arrived at a different degree of credence. Applied to this simple sort of case, the principle separating the two camps amounts to something like this:

Independence: In evaluating the epistemic credentials of another's expressed belief about P, in order to determine how (or whether) to modify my own belief about P, I should do so in a way that doesn't rely on the reasoning behind my initial belief about P.

¹ [Acknowledgments]

² I take the label "Conciliatory" from Elga (forthcoming).

The motivation behind the principle is obvious: it's intended to prevent blatantly question-begging dismissals of the evidence provided by the disagreement of others. It attempts to capture what would be wrong with a P-believer saying, e.g., "Well, so-and-so disagrees with me about P. But since P is true, she's wrong about P. So however reliable she may generally be, I needn't take her disagreement about P as any reason at all to change my belief."

There is clearly something worrisome about this sort of response to the disagreement of others. Used as a general tactic, it would seem to allow a non-expert to dismiss even the disagreement of large numbers of those he took to be experts in the field. And Conciliationism's rejection of this sort of move allows it to deliver intuitively attractive verdicts in many cases involving apparent parity of epistemic credentials. A paradigm example (adapted from Christensen (2007)) is:

Mental Math: After a nice restaurant meal, my friend and I decide to tip 20% and split the check, rounding up to the nearest dollar. As we have done many times, we do the math in our heads. We have long and equally good track records at this (in the cases where we've disagreed, checking with a calculator has shown us right equally frequently); and I have no reason (such as those involving alertness or tiredness, or differential consumption of coffee or wine) for suspecting one of us to be especially good, or bad, at the current reasoning task. I come up with \$43; but then my friend announces that she got \$45.

In such cases, even opponents of Conciliationism concede that I should become much less confident that my share is \$43, and indeed should not be significantly more confident in \$43 than in \$45. Nevertheless, several philosophers have recently offered arguments against Conciliationism in general, aiming to show that the putting aside of one's original reasoning mandated by **Independence** leads to unacceptable consequences in other sorts of cases.

Below, I'll first defend Conciliationism by arguing that **Independence** does not have the unappealing consequences that some have worried about. Having made room for a Conciliationist account, I'll describe some issues that confront the project of developing a full Conciliationist account of rationally responding to disagreement. I'll then argue that these issues must be faced even by reasonable accounts that reject **Independence**. The issues flow from a

certain feature of the wider epistemological territory that has not yet been well explored: rational accommodation of evidence that one has made cognitive errors.

1. Does respecting Independence amount to throwing away evidence?

The first problem I'd like to consider is given forceful development by Thomas Kelly (forthcoming). Kelly argues against what he calls the Equal Weight view: that when I have reason to think my friend is an epistemic peer (that is, that she is generally equally reliable in the domain in question), and have no reason (*independent of my own reasoning about P*) to think her less reliable about P, I should adjust my level of credence in P so as to split the difference with her. His argument proceeds via a series of cases, two of which I will adapt here to my terminology.

Right and Wrong: Right and Wrong are mutually acknowledged peers considering whether P. At t_0 , Right forms 0.2 credence in P, and Wrong forms a 0.8 credence in P. The evidence available to both of them actually supports a 0.2 credence in P.³ Right and Wrong then compare notes, and realize they disagree.

Kelly notes that the Equal Weight view counsels them both to split the difference, each ending up at t_1 with credence 0.5. But this, Kelly argues, is counterintuitive. Before their epistemic compromise, Right and Wrong were in strongly asymmetrical situations. But Kelly says, "For an advocate of the Equal Weight view, this seemingly important asymmetry completely washes out" once Right and Wrong adjust their beliefs:

What is quite implausible, I think, is the suggestion that [Right and Wrong] are rationally required to make *equally* extensive revisions in [their] original opinions, given that [Right's] original opinion was, while [Wrong's] was not, a reasonable response to [their] original evidence. After all, what is reasonable for [them] to believe after [they] meet at

³ For the sake of argument, Kelly grants a principle he thinks false: that evidence will dictate a unique value for rational credence in a proposition.

t1 presumably depends on the total evidence that [they] possess at that point. (Kelly forthcoming, ms. 13)

It seems to me that Kelly is entirely correct in saying that we should not see Right and Wrong as being in epistemically symmetrical situations at t1. To the extent that we did, we'd be overlooking the bearing of the original evidence on what Right and Wrong should believe. And this is in fact the trap Kelly sees the Equal Weight view as falling into: "With respect to playing a role in determining what is reasonable for [them] to believe at t1, [the original evidence] gets completely swamped by purely psychological facts about what [Right and Wrong] believe." (op. cit. 14). The general problem Kelly lays at the feet of the Equal Weight version of Conciliationism, then, is that it makes rational belief in disagreement situations depend completely on the "psychological evidence"—evidence about people's beliefs.

Note that this apparent problem seems to flow from **Independence**: it's Right's inability to justify downgrading Wrong on the basis of Right's own (correct) initial reasoning about P that mandates her compromise. Insofar as Right can resist compromise, she must think that in this case, her initial reasoning was (probably) better than Wrong's. She by hypothesis has no independent reason for thinking this. Thus, her confidence must be rationalized by her reasoning on the matter under dispute.

Before assessing this argument, let us examine a related case that Kelly uses to sharpen his point:

Wrong and Wronger: Wrong and Wronger are mutually acknowledged peers considering whether P. At t0, Wrong forms 0.7 credence in P, and Wronger forms a 0.9 credence in P. The evidence available to both of them actually supports a 0.2 credence in P. Wrong and Wronger then compare notes, and realize they disagree. They follow the dictates of Equal Weight, and at t1 they compromise at 0.8.⁴

⁴ The Equal Weight view may of course be *defined* to require this sort of difference-splitting (and this is indeed a natural and initially appealing form for a Conciliationist view to take). But it's important to see that Conciliationism need not be committed to this general policy. In fact, I would argue that it actually runs counter to the motivating insight behind Conciliationism: that we must take account of the possibility that we've made cognitive mistakes, and that the beliefs of others serve as checks on our cognition. Consider a case where I

Kelly writes:

“On the Equal Weight view, [their] high level of confidence that [P] is true at time₁ is automatically rational, despite the poor job each of [them] has done in evaluating [their] original evidence. ... However, it is dubious that rational belief is so easy to come by.”
(Kelly forthcoming 16)

Again, Kelly’s intuitive verdict on the case seems correct: we should not see Wrong and Wronger’s post-compromise beliefs as rational. Again, to do so would be to treat their original evidence as if it didn’t matter.

Thus it would, I think, be very damaging to Conciliationist views if their insistence on **Independence** amounted to insisting that one’s original evidence was irrelevant to the rationality of the beliefs one ended up with after making one’s conciliatory epistemic compromise. Nevertheless, I think that the most plausible version of the Conciliationist position does not have this consequence.

To see why, let us focus on what Conciliationism is designed to do (for the present, I’ll work with Kelly’s Equal Weight version of Conciliationism). Conciliationism tells us what the proper response is to one particular kind of evidence. Thus the Equal Weight Conciliationist is committed to holding, in Kelly’s cases, that the agents have taken correct account of a particular

come to have fairly high credence—say, .92—in P, as follows: my initial inclination is to be even more certain of P, but I scale back my confidence a bit because I know I make some mistakes. I then learn that my friend, whom I take to be my peer on such matters, has also considered the issue and has become .91 confident in P. I suppose that she arrived at her credence in much the same way as I did. But it seems that learning of her credence should make me more confident that I didn’t make a mistake. If that’s right, I should raise my credence beyond .92, not lower it as difference-splitting would dictate. This verdict is entirely consistent with Independence. It is also consistent with (in the intuitive sense) giving my friend’s opinion equal weight, and even with the view advocated in Elga (2007), whence the term “Equal Weight View” derives. Moreover, there are technical difficulties with the uniform difference-splitting formulation of Conciliationism (see Shogenji (2007) and Jehle and Fitelson (2007)). But while I think that it’s important to note that neither Conciliationism in general, nor giving one’s peer’s opinion equal weight in particular, requires uniform difference-splitting, neither Kelly’s argument, nor my discussion of it, turns on this point.

bit of evidence—the evidence provided by their peer’s disagreement. But having taken correct account of one bit of evidence cannot be equivalent to having beliefs that are rational, all things considered. If one starts out by botching things epistemically, and then takes correct account of one bit of evidence, it’s unlikely that one will end up with fully rational beliefs. And it would surely be asking too much of a principle describing the correct response to peer disagreement to demand that it include a complete recipe for undoing whatever epistemic mistakes one might have made in the past. If Conciliationism is understood in the right way, then, it is not committed to deeming the post-compromise beliefs in **Wrong and Wronger** automatically rational. And in allowing us to criticize Wrong and Wronger’s post-compromise beliefs, Conciliationism thus understood does not entail, or even suggest, that Wrong and Wronger’s original evidence has become irrelevant to the rationality of their post-compromise beliefs.

A similar point applies to the asymmetry in the **Right and Wrong** case. Conciliationism does not entail that Right and Wrong end up with equally rational beliefs. Nor does it entail that they were rationally mandated to make equally extensive revisions to their opinions. Of course, it does have the consequence that the revisions *required by the disagreement* are equally extensive. But this doesn’t erase the fact that Wrong had *other* reasons for revisions, reasons which would mandate greater revisions.

Equal Weight Conciliationism *is* committed to holding that Right’s post-compromise belief is rational (supposing no other background irrationality in the case). But this strikes me as roughly correct.⁵ Right reacted correctly to the original evidence. She then encountered further evidence, which (as it turned out) was misleading. But respecting misleading evidence is no rational defect. So the Conciliationist should be perfectly comfortable with giving her seal of approval to Right’s making major alterations to her original rational belief.

It turns out, then, that Conciliationism’s respecting of **Independence** does not after all render irrelevant the reasoning and evidence on which Conciliatory agents base their initial beliefs.

2. A Follow-Up Objection, and Agent-Specific Evidence

⁵ The reason for the qualification “roughly” will be explained below.

The argument of the previous section shows that Conciliationism doesn't entail throwing away evidence. But the proposed response to Kelly's first case may seem to lay Conciliationism open to a different difficulty. After all, it would seem that Right and Wrong in have exactly the same three bits of evidence:

E1: The original evidence relevant to P.

E2: The fact that Right reached credence 0.2 on the basis of E1.

E3: The fact that Wrong reached credence 0.8 on the basis of E1.

And if each of them has in the end reached the same credence on the basis of the same evidence, how can we say that Wrong's credence falls short rationally, while Right's does not?⁶

To begin thinking about this puzzle, suppose we approached the example by considering a third party confronted with E1-E3, and asking what, from a Conciliationist point of view, she should believe about P. In fact, it seems to me that such an agent confronted with E1-E3 should not end up giving P 0.5 credence, as I've claimed that the über-rational Right should. Such an agent should of course take the import of E1 to be to rationalize 0.2 credence in P. But then she'd see that one other agent agrees and one disagrees. The undercutting power of Wrong's belief is diluted by the supporting power of Right's. So the agent should end up with credence somewhere below 0.5.⁷ And this is in fact where Kelly thinks Right should end up.

But does our conclusion about the third party carry over to Right? Interestingly, I think it does not. And the reason for this involves a facet of the epistemology of disagreement that hasn't been fully articulated: that the evidential force of the information expressed in claims like E2 and E3 depends crucially on whether the agent responding to the evidence is identical to one

⁶ This way of describing the evidence is from Kelly (forthcoming). The puzzle is also due to Kelly, in conversation.

⁷ This assumes that the agent has some reason for epistemically respecting Right and Wrong, so she should, from a Conciliationist viewpoint, take their views into account. It also assumes that the agent does not see Right and Wrong as such experts (relative to herself) that she should not even try to figure out the import of the evidence directly, but instead should just base her beliefs on Right's and Wrong's. One way of avoiding both issues: stipulate that the agent has excellent evidence of peerhood with Right and Wrong.

of agents mentioned in E2 and E3. In order to see how this dependence works, let us first consider a simpler case involving evidence of possible cognitive malfunction.

Suppose I'm participating in placebo-controlled trials of a reason-distorting drug. The drug has been shown to cause people to make mistakes in algebraic reasoning, but to leave most of their cognitive faculties unscathed. Moreover, those affected by the drug do not notice that their algebraic reasoning is impaired; in fact, they seem to themselves to be thinking as clearly and distinctly as they ever do. I've been through several trials, some with the drug and some with the placebo, and I've never seemed to myself to have been affected; but watching the tapes of myself in previous trials, I see myself earnestly—even heatedly—insisting on the patently mistaken conclusions I've drawn on the assigned problems when I got the active pills. It seems clear that, in such a situation, if I'm given a pill and then asked to draw a conclusion from some evidence that requires algebraic interpretation, I should be far less confident of my answer than I ordinarily would be.

Now suppose we represented my evidence as follows:

E1: The evidence presented as part of the experimental problem

E2: [NN] had a 50% chance of taking an active pill

Clearly, a rational third party presented with E1 and E2 would not be much bothered by E2. In fact, E2 seems like it should be completely evidentially irrelevant to the belief one should end up with about the algebra problem—for everyone except [NN]. (More precisely, E2's relevance for an agent will depend on the degree to which the agent believes she is [NN]. One might even want to factor in whether the agent's confidence that she's [NN] is rational. But let us leave these complications aside, and just consider cases where agents have rational and correct beliefs about their identities.) So: in this sort of case, the rational import of evidence is *agent-specific*.

Now when I'm confident that P, and find out that my friend is confident that not-P, the evidence provided by disagreement is at least partly of a similar sort. Given that my friend and I have access to the same first-order evidence, her disagreement with me is also evidence that I've misconstrued the import of that first-order evidence. In this respect, it's like the undermining evidence in the drug example. This suggests that the evidence Right and Wrong have in the case above might be described more perspicuously as follows:

Right's evidence is:

E1: The original evidence relevant to P.

E2r: The fact that I reached my present credence 0.2 on the basis of E1.

E3r: The fact that my peer reached credence 0.8 on the basis of E1.

Wrong's evidence is:

E1: The original evidence relevant to P.

E2w: The fact that my peer reached credence 0.2 on the basis of E1.

E3w: The fact that I reached my present credence 0.8 on the basis of E1.

Now how should the difference between Right's and Wrong's evidential situations affect their respective credences in P? The answer to this question depends on how the identity of the agents figures into the epistemic import of the bits of evidence described above.

Consider first how an agent should regard the information that she herself has reached a certain conclusion from her evidence. Suppose I do some calculations in my head, and become reasonably confident of the answer 43. I then reflect on the fact that I just got 43. It does not seem that this reflection should occasion any change in my confidence. On the other hand, suppose I learn that my reliable friend got 43. This, it seems, should make me more confident in my answer. Similarly, if I learn that my friend got 45, this should make me less confident.⁸

The fact that the first-person psychological evidence is relatively inert in this respect is exactly what one would expect, given the main intuitive rationale for adjusting one's beliefs in the face of disagreement with an equally-informed friend. Since I recognize that I may sometimes misconstrue the import of evidence, I see that my friend's reaction to the same evidence may well confirm or disconfirm my having assessed that evidence correctly. But clearly, I cannot use my own reaction to the evidence as a check in this way. Thus for me, psychological reports about others serve as a kind of epistemic resource that psychological reports about myself do not serve.

⁷ I do not expect these judgment to be controversial. Even non-Conciliationist philosophers concede that in cases like this, the disagreement of a friend should make me less confident; and taking agreement of a friend to justify increased confidence is just the other side of the same coin.

There is a sense, then, in which Right and Wrong have different evidence to react to.⁹ In each case, we may take the first-person psychological evidence to be incapable of providing the sort of check on one's reasoning that third-person evidence provides. In this sense, it is relatively inert. So the important determinants of what's rational for Right to believe are the original evidence E1 (which should, and does, move her to put 0.2 credence in P), and Wrong's dissent (which does and, according to the Equal Weight Conciliationist, should move her from 0.2 to 0.5). In contrast, the determinants of what Wrong should believe are E1 (which should move him toward having 0.2 credence in P), and Right's belief (which *also* should move him toward 0.2). Looked at this way, it's not surprising that his arriving at 0.5 rather than 0.2 is less than fully rational.

The upshot is this: Right's and Wrong's evidential situations are not symmetrical. Upon closer examination, it turns out that their two situations do not call for the same degree of confidence in P. And thus when Right and Wrong arrive at the same degree of confidence in P, the Conciliationist need not consider their degrees of confidence equally rational, or equally supported by the evidence.

Understanding the power of disagreement-based evidence in this way also disarms a related worry about the Equal Weight version of Conciliationism voiced in Kelly (forthcoming). Suppose we grant that the correct response to disagreement is not completely Steadfast—that peer disagreement should typically occasion *some* change of belief. Still, in a case like **Right and Wrong**, Kelly notes that the psychological evidence is *balanced*: Right's belief points toward not-P just as strongly as Wrong's belief points toward P. Such balanced psychological evidence “tends to push what it is reasonable for us to believe about the hypothesis in the direction of agnosticism” (forthcoming, ms 33). Thus, given that the non-psychological evidence strongly favors not-P, it's reasonable to expect that the total evidence in the example favors not-P, though less strongly than does the non-psychological evidence alone. If this “balancing argument” is right, then even if we admit some conciliation, the correct credence in P to adopt here would seem to fall well below Right's 0.5, *contra* Equal Weight conciliationism.

⁹ I'm not sure that it's quite right to say that they have different evidence, rather than that their different positions make the rational import of their common evidence different. I don't think anything important hangs on this.

As we saw above, this verdict is in a way exactly correct: it nicely describes how a third party should evaluate E1 – E3. But if we describe the case in a way that abstracts from whether the person confronting the evidence is a third party, or is one of the subjects of the psychological evidence, we will miss an important determinant of rational belief. The proponent of Equal Weight Conciliationism should concur in Kelly’s verdict on a third-party version of the example. But she should dissent if the description is meant to apply to Right’s beliefs. For Right, if she takes account of the total evidence as she ought to, will take psychological information about her friend’s beliefs to be important evidence, in a way that psychological information about her own beliefs is not. For her, then, the balancing argument does not apply.¹⁰

In sum: A strongly Conciliationist view is perfectly consistent with our judgments about rational responses to the total evidence in cases like **Right and Wrong**. In fact, it helps connect the epistemology of disagreement to a more general epistemic phenomenon: the special way in which evidence of a certain agent’s possible cognitive malfunction should inform that particular agent’s beliefs.

3. Hard Cases: Extremely High Rational Confidence

I would like in the next few sections to turn to examine quite a different sort worry about Conciliationism: that it gets certain cases clearly wrong. The first sort of hard cases are ones

¹⁰ Analogues of the balancing argument clearly fail in cases of not related to disagreement. Suppose that Jocko wanders into an art museum, and beholds *Study: 4b*, the first painting he comes to. It appears to him to be a simple red rectangle. Jocko concludes that (a) *Study: 4b* is a red rectangle, and (b) the museum’s current exhibit is unlikely to prove rewarding for him. As he leaves, he notices an artist’s statement explaining the show. The artist has painted 50% of the canvasses red, and 50% white, and then lit the white ones with deceptive lighting so that they look just like the red ones. Considering Jocko’s total evidence as to *Study: 4b*’s redness, we now have:

- E1: the appearance of *Study: 4b*
- E2: the information on the sign

Clearly, when Jocko had just E1 to go on, he was reasonable in coming to believe that *Study: 4b* was red; E1 by itself strongly favors this conclusion. And E2 is balanced, in the sense that it pushes the rational believer towards agnosticism regarding *Study: 4b*’s redness. But it clearly does not follow that the total evidence E1 and E2 favors having greater than 0.5 credence that *Study: 4b* is red.

where an agent begins with extremely high rational confidence in her belief. In various such cases, it seems wrong to hold that she should revise her belief much at all, even if the agent's friend disagrees sharply, and even if, before discovering the disagreement, she would have considered the friend her epistemic peer on the sort of issue in question. This suggests that it is, after all, legitimate for the agent to demote her friend's dissenting opinion on the basis of her own reasoning on the matter under dispute. In other words, it suggests that **Independence** fails in these cases.

Let me begin with an example based on similar examples in forthcoming papers by Jennifer Lackey, Ernest Sosa, and Bryan Frances:

Careful Checking: I consider my friend my peer on matters of simple math. She and I are in a restaurant, figuring our shares of the bill plus 20% tip, rounded up to the nearest dollar. The total on the bill is clearly visible in unambiguous numbers. Instead of doing the math once in my head, I take out a pencil and paper and carefully go through the problem. I then carefully check my answer, and it checks out. I then take out my well-tested calculator, and redo the problem and check the result in a few different ways. As I do all of this I feel fully clear and alert. Each time I do the problem, I get the exact same answer, \$43, and each time I check this answer, it checks out correctly. Since the math problem is so easy, and I've calculated and checked my answer so carefully in several independent ways, I now have an extremely high degree of rational confidence that our shares are \$43. Then something very strange happens. My friend announces that she got \$45! Here, many people feel that I should not reduce my confidence in \$43 very far at all. And this intuition holds even if we stipulate that I could see my friend writing numbers on paper and pushing calculator buttons, and that my friend assures me that she did her calculations slowly and carefully, felt clear while doing them, and got her same answer repeatedly. It seems that I'd be reasonable in this case to suspect strongly that something screwy must be going on with my friend.¹¹

¹¹ Sosa and Lackey also discuss somewhat more extreme, but less realistic, versions of this type of example: disagreement about maximally clear perceptual belief (e.g., when I see someone sitting at the table with us, and my friend claims that there's no one there), and disagreement about elementary math (e.g. my friend insists that $2+2=5$). I worry a bit about intuitions based on such far-fetched examples. Nevertheless, I think that the discussion below of **Careful Checking** will apply to them as well.

This intuition—which to a large degree I share—seems to cut directly against Conciliationism, and particularly against **Independence**. Why, after all, do I suspect that something screwy went on with my friend? It’s just because she reported getting \$45. And the only reason that that would indicate anything amiss is that I’m quite sure that the answer is not \$45. Yet my reason for being so sure that the answer is not \$45 is just my own meticulous reasoning showing it to be \$43! Thus, in describing a similar case, Sosa writes: “Now I am in the Moore-like position of having to say that if his procedure has led to that result, there must be something wrong with his procedure. ... I still lack *independent* reason to downgrade my opponent’s relevant judgment and his epistemic credentials on the question that divides us. Only based on our disagreement can I now demote him.” (forthcoming, ms. 18-19). And Lackey says that cases involving extremely high justified confidence “show precisely why condition (1) [a formulation of **Independence** from Christensen (2007)] should be eliminated from Christensen’s account” (forthcoming a, ms. P. 45, fn. 33).

I think that, on closer inspection, cases involving ultra-high initial rational confidence do not end up undermining **Independence**.¹² Let us begin by considering what I should think of my initial opinion in **Careful Checking**. Being generally competent at elementary math problems, having done the calculations repeatedly and carefully both on paper and with a well-tested calculator, having checked the answer in multiple independent ways, and feeling very clear-headed and alert throughout, I should think that it would be *extremely* unlikely for someone in my situation to have gotten (and verified) the same wrong answer each time. That goes hand-in-hand with the legitimacy of my having ultra-high confidence in my answer.

But if that’s right, here’s something else that would be extremely unlikely: two people, both generally competent at elementary math, who worked on the same problem, each having done the calculations repeatedly and carefully both on paper and with a well-tested calculator, each having checked the answer in multiple independent ways, each feeling very clear-headed and alert throughout, and each repeatedly coming up with (and verifying) a different answer. This is important, because it means that, in the strange scenario described in **Careful Checking**, I have good reason to think that something screwy has gone on.

¹² My explanation for this fills out the strategy briefly sketched in Christensen (2007, 200-203). It also draws heavily on Lackey’s (forthcoming a, b) insightful analysis of this type of example. The conclusion I’ll draw about **Independence**, however, is opposite from hers.

What possible explanations are there for the divergence between our announced answers in **Careful Checking**? Well, one possibility is that one of us has experienced some bizarre mental malfunction resulting in errors that somehow led to the same wrong answer in all the independent ways of doing and checking the problem. Another is that one of us is actually exhausted, or drunk, or tripping, or experiencing a confusing psychotic episode, and is really only managing to go through the external motions of recalculating and checking, without actually paying clear attention. Still another is that one of us is just joking, or messing with the other's head for fun. Another is that one of us is deliberately making false claims about his or her answer, for the pure thrill of bald-faced lying, or as part of a psychological or philosophical experiment, or perhaps in an earnest attempt to problematize the hegemony of phallogocentric objectivity by an act of performance art.

This is not an exhaustive list of explanations for the divergence of our announced answers. But it is enough, I think, to show why I am in a position to think that the answer my friend announced is less likely to be correct than mine is. For example, while I can definitively rule out the possibility that I've deliberately announced an incorrect answer for recreational, experimental or performance-artistic reasons, I cannot be nearly so sure of ruling out these possibilities for my friend. Similarly, while I can be very sure that I was actually paying attention rather than going through the motions of checking my answer, I cannot be nearly so sure that my friend was. And while there are *conceivable* sorts of mental malfunction that would affect my reasoning without my having any sign of trouble, most reason-distorting mental malfunctions come with clear indications of possible trouble: dizziness, seeing patterns moving on the wall, memories of recent drug-taking or of psychotic episodes. And I'm in a much better position to rule these out for myself than I am for my friend. Let me put the information I'm depending on in all these cases under the common label, taken from Lackey, of "personal information." The personal information I have about myself in **Careful Checking** provides a perfectly reasonable basis for my continuing to think that our shares of the bill are much more likely to be \$43 than \$45, despite discovering the disagreement of my heretofore-equally-reliable friend.¹³

¹³ My usage of "personal information" here encompasses a wider range of examples than those mentioned by Lackey. She includes information relevant to various possibilities of cognitive malfunction; I've extended it to include information relevant to possibilities of

If this is my basis for maintaining my belief, have I violated **Independence**? It seems to me that I have not. True, in supporting my suspicion that something screwy has gone on with my friend, I relied on the claim that I arrived at my answer to the math problem by a very reliable method. But my reasoning did not rely on the *results* of my calculations at all. I did *not* say, “Well I’m very sure the answer is \$43. My friend says it’s \$45, so something screwy must have gone on with her.” That sort of reasoning would indeed violate **Independence**. But the reasoning I used was quite different. It was more like this: “I arrived at my answer by an extremely reliable method. It is very unlikely that two people employing such a method would end up sincerely announcing incompatible beliefs. The belief my friend announced was incompatible with the one at which I arrived. So it’s likely that one of us did not arrive at his or her belief in the highly reliable way, or that one of us is not sincerely announcing his or her beliefs. I can eliminate (via personal information) many of the ways that I could have failed to use the method, as well as the possibility that my announcement was not sincere. But I cannot eliminate analogous possibilities for my friend. So it’s likely that she did not sincerely announce a belief that was formed by the highly reliable method.”

Notice that this reasoning does not even refer to the particular answer I got. In fact, the reasoning could have been formulated in advance of my doing any calculation, or even seeing the bill! This shows that while the reasoning relies on certain facts *about the kind of reasoning I use*, it does not rely on my reasoning itself. It takes into account the *fact* that we disagreed, but it does not depend on the substance of the disagreement. So it does not beg the question against my friend’s belief in the way **Independence** is designed to prevent.

Thus my reason for maintaining my belief in this case is entirely consonant with the sort of positions advocated by Conciliationist writers. It is obvious that in considering the epistemic import of one’s friend’s expressed beliefs, one must take into account certain facts about one’s reasoning. If I know that I’ve been reasoning while tripping, or if I know that the reasoning

insincere assertion. But I think that my usage of the phrase, as well as the role I give personal information in assessing the evidential force of my friend’s disagreement, is very much in the spirit of Lackey’s analysis.

Frances (forthcoming) and Fumerton (forthcoming) also argue that in certain disagreements where one begins with ultra-high rational confidence, one will reasonably suspect that one’s friend is joking or crazy, and thus one needn’t revise one’s belief. Neither makes this point in the context of evaluating Independence, (though in other parts of their papers, Frances seems sympathetic to, and Fumerton seems to deny, something like Independence).

method I've used is only moderately reliable, that gives me reason to accord more weight to my friend's disagreement (to the extent, of course, that I doubt that her own reasoning suffers from these sorts of weaknesses). Similarly, to the extent that I know my own reasoning to have been of a particularly reliable sort, that gives me reason to give my friend's disagreement less weight (to the extent, again, that I doubt that her own reasoning is of this same particularly reliable sort). And I may bring these sorts of considerations to bear without relying on my own initial reasoning concerning the disputed matter.

It seems to me that this sort of treatment applies particularly nicely to the above-mentioned extreme examples offered by critics of **Independence**: cases where one's friend claims to believe that there's no one else at the table, or that $2+2=5$. If such a bizarre situation were actually to occur, I think one would reasonably take it as extremely unlikely that one's friend (a) was feeling as clear-headed as oneself; (b) had no memories of recent drug-ingestions or psychotic episodes; and, most importantly, (c) was being completely sincere. Thus, to use Lackey's term, one's personal information (that one was feeling clear, lacked memories suggesting mental malfunction, and was being sincere in one's assertion) would introduce a relevant asymmetry, and one could reasonably maintain one's belief. Even the single possibility that my friend was obnoxiously messing with my head, in part precisely by assuring me repeatedly that all was clear and sincere on her end, would be *far* more likely than the possibility that the two of us were engaged in a sober and sincere disagreement over whether there was another friend at the table, or whether $2+2$ added up to 4. But nothing in this reasoning undermines **Independence**.

4. An Objection to this Analysis

It is worth considering one obvious objection to the claim that maintaining belief in the above cases is consistent with **Independence**. The objection is based on a comparison between the cases we've been discussing and cases involving significantly lower degrees of initial rational confidence. One might point out that the personal information, which provides the independent basis for my thinking myself more likely to be right in **Careful Checking**, is also present in **Mental Math**, where one clearly should reduce one's confidence dramatically. And as Lackey points out, the obvious difference between this case and **Careful Checking** is simply the degree of rational confidence I have in my initial opinion; so in some way, my high initial

rational confidence enables the personal information to play its key role in **Careful Checking**. This might lead one to suspect, then, that my maintaining my belief in **Careful Checking** must after all rely on my reasoning concerning the disputed matter.

I think, though, that a close look at how rational confidence and the efficacy of personal information are related reveals that this is not so. Consider how one would explain my friend's expressed disagreement in **Mental Math**. I know that doing a problem once in my head is not an extremely reliable process, because people commonly make undetected slips in mental calculation. So the overwhelmingly likely explanation for our disagreement obviously lies in one of us making this everyday sort of slip. Unfortunately, my personal information does not help me to eliminate this possibility for myself. Of course, there are also the exotic possibilities considered above: that one of us is drunk, tripping, psychotic, joking, lying or engaging in performance art. And my personal information does allow me to eliminate various exotic possibilities for myself, and not for my friend. But since these exotic scenarios are so unlikely, the fact that I can eliminate some of them has only a tiny effect on the plausibility of explaining the disagreement in a way that involves the falsity of my friend's claim. That is why I should (in categorical terms) suspend belief, or (in graded terms) come close to splitting the difference with my friend, in the sense of seeing the two answers as about equally likely to be correct.¹⁴

In **Careful Checking**, by contrast, the high degree of rational confidence I have in my initial belief is correlated with my rationally taking my reasoning method to be extremely reliable. And it is the extreme reliability of this method, a method which eliminates the "everyday mental slip" explanation of our disagreement, which both makes this sort of disagreement so unusual, and makes the exotic explanations vastly more probable, should a disagreement occur. (This is why it's only in these cases that I'll think that something screwy must be going on.) At this point, when personal information allows me to eliminate several exotic possibilities for myself, but not for my friend, the balance of probability is shifted dramatically over to explanations involving the falsity of my friend's expressed belief.

Thus it turns out that my high degree of initial rational confidence is correlated with my legitimately maintaining my belief in certain cases. It's correlated because high initial

¹⁴ It's worth noting that even eliminating a few highly improbable exotic scenarios allows me to favor my own belief a tiny bit. So the availability of personal information does mean that I should not *exactly* split the difference, even in **Mental Math**. But Conciliationism should not be seen as saying otherwise.

confidence is appropriate only when one may rationally take one's reasoning method to be extremely reliable, which in turn eliminates everyday explanations for the disagreement, and makes exotic explanations—which tend to be sensitive to personal information—much more probable. But none of this undermines **Independence**. For in adjudicating explanations for our disagreement in any of these cases, I do not rely on my reasoning about the disputed matter.

Before leaving discussion of cases involving extremely high rational confidence, it's worth emphasizing a point about how these cases relate to Conciliationism in general. It's obvious that most of the issues that are subject to controversy are nothing like the issue of whether our friend is before us, or whether $2+2=4$. A hallmark of the latter cases—which is intimately related to their involving extremely high levels of rational confidence—is exactly that beliefs formed in these ways are virtually never subject to disagreement. So it's worth noting that even if the Conciliationist shares the Steadfast view's verdict on cases involving extremely high rational confidence, there is no reason to think that the rational permissibility of maintaining one's belief in these cases will bleed over into the controversial cases which give the disagreement issue some of its urgency. This is the reason that the theoretical diagnosis of the extreme cases—and in particular, the question of whether they require violation of **Independence**—is important.

5. Hard Cases, Cont'd: Multiple Disagreements

There is another kind of case which puts at least prima facie pressure on Conciliationism. Let me illustrate it with an example based on one from Kelly (forthcoming):

Seminar: I'm in a meteorology graduate seminar with Stranger, another graduate student. I don't know him very well, but his first few comments seem quite sensible to me. I take myself to be a pretty reliable thinker in meteorology, though not more reliable than most grad students. At the break, I discover that we've both read a fair amount about issue P, but while I'm quite confident that P, Stranger expresses equal confidence that $\sim P$. I, being a good Conciliationist, then become significantly less confident in P. But as the conversation develops, I find that Stranger and I disagree equally sharply about Q, R, S, T and so on—a huge list of claims. And these claims are not part of some tightly interconnected set of claims that would be expected to stand or fall together: they're largely independent of one another. Do I now have to become

significantly less confident about all of them? Here, it might well seem intuitively more reasonable for me to stop putting so much stock in Stranger's claims. As Kelly notes, (forthcoming ms. p. 54), it seems that I should instead reevaluate my original opinion of Stranger, and become increasingly confident that I'm better at evaluating evidence in this field. In fact, it seems that the reasonable response to the repeated disagreements will include moving back to being quite confident in P.¹⁵

This might seem to cut against Conciliationism. After all, if Conciliationism requires me to become much less confident in P when we disagree about P, it might seem that it cannot then allow me to use repeated disagreements between us asymmetrically to lower my general trust in Stranger's beliefs and regain my confidence in P.

It turns out, though, that Conciliationism does allow exactly this to happen. This is shown by the following (fairly realistic) filling out of the **Seminar** case:

Seminar, Cont'd: In addition to believing antecedently that I'm quite reliable in meteorology, and that the vast majority of others are about equally reliable, I also believe that there are a very few people—call them meteorologically deranged—who are horribly unreliable. I'm extremely confident that I'm not one of them. And I take such people to be rare enough that my estimation of my own reliability is not much different from my estimation of the reliability of a random person in the field.

Given these assumptions, when I first find out that Stranger and I disagree about P, Conciliationism would counsel me to become much less confident in P. But when we discuss 25 more claims, and he disagrees with me about all of them, I should now think: given the extent of our disagreements, it's incredibly unlikely that Stranger and I are both very reliable at all. The more disagreements we discover, the more likely it is that one of us is deranged. Since I'm more confident (independent of the disagreement) that I'm not deranged than that Stranger isn't, I

¹⁵ The structure of the example is Kelly's, though I've filled in or changed various aspects of the case. Kelly presses this example as a counterargument to Elga's(2007) bootstrapping argument for his Equal Weight view. Kelly also specifies in his example that as a matter of fact, my initial opinion on the issues under dispute is in fact the rational one. But it seems to me that even without making this assumption, the example elicits the intuitions in question.

should become more confident that I'm better at evaluating the evidence than Stranger is. And so I need not become much less confident about all the things we disagree about. Thus, even on the Conciliationist view, I need not be driven to widespread agnosticism by Stranger's repeated disagreement. And indeed, I should regain most of my original confidence in P.

It's important to see that my reevaluation of Stranger here is entirely consistent with **Independence**. I'm not using my beliefs that P, Q, R, etc. as premises to show that he's wrong about many things, and hence is unreliable. I'm just using facts about our reasoning: that the wide extent of our disagreement indicates that one of us is seriously malfunctioning epistemically. Of course, the end result does depend on an initial asymmetry between my assessment of myself and my assessment of others: I'm more confident that I'm not deranged than I am that an arbitrary other person is not deranged. But the example shows that this is quite consistent with my taking others, about whom I know little, to be about as likely as I am to get particular claims right. It is this latter attitude which is behind Conciliationism's recommendations in many cases to suspend belief on learning of a particular disagreement.

The way I've expressed the agent's attitudes in the case above distinguishes between "single claims" and large conjunctions of claims. It's worth noting that this convenient classification need not bear heavy theoretical weight. To bring this point out, suppose someone objected as follows to the above analysis: "You say that if Stranger disagrees about just one claim P, you should become agnostic on P. But suppose that Stranger's *initial* claim is ($\sim P \ \& \ \sim Q \ \& \ \dots \ \& \ \sim Z$), where the conjuncts are the negations of all the particular claims involved in the repeated disagreement you described above. Does Conciliationism now say you should give significant credence to *this* claim? That would mean taking it as reasonably likely that all of P, Q, R, etc. are false, which would require your becoming agnostic (at best) about each individual claim."

To answer this question, we should note that it's no part of Conciliationism that one take similar attitudes to all of the propositions one believes. I may, as stipulated in **Seminar**, have fairly high confidence in each of P through Z. But there are other claims I have much greater confidence in. Given my fairly high confidence in P through Z, and given their relative independence from one another, I ought to have extremely high confidence that they're not all false, i.e. that $\sim(\sim P \ \& \ \sim Q \ \& \ \dots \ \sim Z)$. This is just another way of saying that I'm extremely highly confident that I'm not horribly screwed up epistemically in meteorology. So if Stranger had

asserted his big conjunction right at the beginning of the conversation, Conciliationism would not have required me to suspend belief.

It's clear that there's nothing unusual about having the sort of distribution of confidence among the sort of "individual" and "conjunctive" claims involved in **Seminar**. In fact, it would seem quite the typical case. So it seems to me that Conciliationism can give quite a nice account of the contrast between disagreement about an isolated particular claims, and disagreement about multiple (or highly conjunctive) claims.

In case the reader is suspicious about the structure of the **Seminar** example, here is a model with the same structure that abstracts from any controversies surrounding disagreement: Consider a situation in which each person has an almanac, and each person's almanac is unique (they're not copies of the same book). Suppose that I have extremely good reason to believe that my almanac is highly accurate—that 98% of the time it's correct. Moreover, I have good reason to believe that the vast majority of other almanacs are as reliable as mine. But I also believe that there are a small number of horribly inaccurate ones—perhaps the Acme company produced them without researching any facts. (This corresponds, in an artificially simplified way, to the sort distribution of reliability among people supposed in **Seminar**.) So I take the distribution of accuracy in almanacs to be this:

99% are 98% accurate

1% are 1% accurate

On these assumptions, the accuracy of my almanac is close to the average accuracy of almanacs in the population at large (the average accuracy is a bit over 97%).

In this situation, if my almanac says P, I should initially be very confident that P is true. But if I then find out that a Stranger's almanac, about which I have no special information, disagrees, my confidence in P should drop dramatically. This is because in single-claim matchups between my almanac and a randomly selected one, I should expect mine to be wrong in a fair proportion of the (rare) cases when they disagree. However, if I look up 16 more facts at random in my almanac, and the Stranger's almanac disagrees on every one of these claims, I should not dramatically lower my confidence in all 17 claims. Instead, I should conclude that the Stranger's almanac is very likely an Acme, and in fact I should then go back to trusting what my

almanac says about even the first of the 17 claims. This is because, in matchups between two almanacs, when they disagree on every one of 17 claims looked up, the overwhelmingly likely situation is that one of them is an Acme (although Acmes are somewhat rare, the chance of two highly reliable almanacs disagreeing 17 times in a row is incredibly miniscule). And we've stipulated that I'm initially extremely sure that mine's not an Acme. This initial asymmetry— analogous to my being much more confident that I'm not meteorologically deranged than that the Stranger isn't—is all that's needed to produce the Conciliatory structure described above.^{16,17}

In sum, then, it turns out that Conciliationism can easily accommodate the intuition that in certain cases, repeated disagreements with a stranger, particularly about relatively independent claims, should occasion little change in belief, while a single disagreement should significantly reduce my confidence. And it seems to me that this is not a mere formal possibility. In many cases, the background assumptions needed to underlie this phenomenon are reasonable: one believes that oneself, and people generally, are pretty reliable about a certain field; one believes that there are a few people who are epistemically very screwed up and thus terribly unreliable; and one is extremely confident that one is not one of those people.

¹⁶ Various assumptions would be needed for a precise version of this example, e.g. the degree to which mistakes in different almanacs are independent of one another, or the probability that when mistakes are made, different almanacs will make the same false claim. But the example seems clear enough without getting into deep technical details.

¹⁷ In some minor variants of this example, where I'm sure that my almanac is slightly less reliable than the vast majority, I should think my almanac and the Stranger's to be exactly equally likely to have gotten P right after the first disagreement. One might wonder how it could be that I should do this, yet regain my confidence in P once the subsequent disagreements arise. It might seem that, if I should think after the first disagreement that my almanac and Stranger's are equally likely to have gotten P right, then this disagreement hasn't favored my almanac's accuracy over the Stranger's, so we're back to square one. But this isn't quite right. Even though I should think the two almanacs are equally likely to be right *about P*, the disagreement about P should also significantly reduce how accurate I should expect the Stranger's almanac to be *in general*. That's because if his is wrong in this case—which is about as likely as not—it stands a fair chance of being a horribly inaccurate Acme (in the long run, about 1/3 of false claims are made by Acmes). On the other hand, if mine is wrong about P, it's still a 98% accurate almanac which got P wrong.

6. Does Independence Lead to Wholesale Skepticism?¹⁸

One might think that the analysis given in the previous section of responses to apparently deranged strangers was too easy, by the Conciliationist's own standards. After all, **Independence** requires me to abstain from relying on the reasoning under dispute in evaluating the epistemic credentials of another. But it's clearly possible for a certain sort of Stranger to dispute a very wide swath of my beliefs—a swath that includes the claims I was relying on in the previous analysis. Suppose that a Stranger, after expressing disagreement about P, adds, "Most people are completely screwed up. Only a few of us are sane. There's massive delusion about." Suppose he's wearing an aluminum foil hat, and he tells me he's just seen Elvis's image in his morning waffle, and that Elvis told him that only those with foil hats can think straight at all. Such a Stranger is denying a claim (that I'm not incredibly screwed up) confidence in which underlay my response to repeated disagreement in the last section. And this, combined with **Independence**, seems to mean that I can't dismiss his wholesale challenge to my beliefs by relying on my confidence in this claim. On the other hand, it's also intuitively obvious that I shouldn't be too worried by this guy's opinion. But given Conciliationists' adherence to **Independence** principles, how can they resist recommending extensive belief-revision in this sort of case?

We should start by noting that **Independence**, by itself, does not mandate anything about belief-change. It tells us to evaluate the epistemic credentials of the other person's belief in a dispute-independent manner. But it doesn't say what to do with the evaluation. To get a Conciliationist result, one must add a principle mandating belief-revision in the event that the evaluation yields certain results. But this sort of principle might be drawn up in different ways. Let us confine discussion at this point to cases where another person disagrees sharply with my initial belief.

One way of drawing up a belief-revision principle would be:

¹⁸ The problem considered in this section has not to my knowledge been raised in print. Versions of it have come up in discussions with several people, but were pressed on me particularly effectively by Ernie Sosa and Josh Schechter.

(A) Insofar as the dispute-independent evaluation *fails to give me* good reason for confidence that I'm better informed, or more likely to have reasoned from the evidence correctly, I must revise my belief in the direction of the other person's.

Another is:

(B) Insofar as the dispute-independent evaluation *gives me* good reason to be confident that the other person is equally well-informed, and equally likely to have reasoned from the evidence correctly, I must revise my belief in the direction of the other person's.¹⁹

Clearly, the first sort of principle threatens to lead straight to skepticism. For suppose that the other person disagrees with virtually all my beliefs about the world. That would leave me with virtually nothing on which to base a dispute-independent evaluation of the relative likelihood of his reasoning correctly. But such a baseless evaluation would clearly not give me a good reason for confidence in my having reasoned correctly; so principle (A) would require massive revision in my beliefs. The problem with (A) is that it in effect turns out to require that one have a non-question-begging response to the skeptic.

On the other hand, the second sort of principle does not obviously have the same vulnerability. For if the other person challenges virtually all of my beliefs, an evaluation based on dispute-independent grounds—i.e., an evaluation based on virtually nothing—cannot provide good reason for me to be confident about either party's likelihood of having reasoned correctly.

On the second sort of view, when disagreement undermines one's rational confidence in some claim, the undermining must be based on one's beliefs about the other person. When those beliefs include extensive dispute-independent evidence of intellectual and evidential parity (as in the **Mental Math** case), the undermining power of disagreement is high. But in some cases, one has little dispute-independent reason to be highly confident, one way or the other, about whether the other person is even one's approximate peer. In those cases, the undermining power of disagreement should intuitively be less.

¹⁹ The contrast between (A) and (B) bears strong similarities to that between Harman's (1986) Principles of Negative Undermining and Positive Undermining.

Thus the “insofar” in principle (B) should be understood as indicating that the undermining power of disagreement by apparent epistemic equals is not all-or-nothing. When one evaluates the epistemic credentials of another’s beliefs in a dispute-neutral manner, one may get varying strengths of reasons for thinking that the other person is as well-informed and as likely to have reasoned correctly as one has oneself. The stronger one’s reasons for thinking equally well of the other’s epistemic credentials, the more one should revise one’s beliefs. And this, I think, is how it should be intuitively.

I suggest, then, that the Conciliatory position should be understood as involving a principle along the lines of (B). I should note, though, that (B) is far from being a precise or sufficiently general principle to provide a recipe for reacting to disagreement evidence. (Some of the reasons for this will be explored below.) But something along these lines would seem to be a promising way of allowing us to respect the anti-question-begging motivations for **Independence**, without sliding into skepticism.

This sort of approach may also help Conciliationism mesh with another intuitive judgment some have defended. It has been suggested that when the proposition in dispute is part of a deep network of beliefs, all of which are challenged by the other person, the rational pressure to revise beliefs should be less.²⁰ Elga (2007, 495 ff.) considers the case of Ann and Beth, who disagree not only over the morality of abortion, but about a whole nest of related moral, religious and ordinary factual matters. Elga suggests that in this sort of case, there will be no fact of the matter about Ann’s opinion, independent of the issues under dispute, of Beth’s likelihood of being correct about abortion. I’m not sure this is right—after all, there may be quite a bit that Ann does know about Beth—that she’s human, intelligent, familiar with certain arguments, honest, etc. This might well yield some dispute-independent assessment of Beth’s reliability.²¹ But even putting this worry aside, the phenomenon in question is intuitively a gradual one: as the area of disagreement becomes greater, the pressure for conciliation diminishes. But it is hard to see how the transition from there being a fact of the matter about Ann’s dispute-independent opinion of Beth to there being no fact of the matter would be able to account for this gradual phenomenon.

²⁰ See Pettit (2006) and Elga (2007).

²¹ See Kornblith (forthcoming) for an argument along these lines.

By contrast, on a (B)-style view, as the area of disagreement gets wider, the basis for dispute-independent assessment gets narrower, and thus the strength of the reasons provided by that narrower basis will often be reduced, and in a gradual manner. So insofar as we find it plausible that the rational pressure to revise one's beliefs diminishes as the extent of disagreement with the other person increases, a principle along the lines of (B) is an attractive option.

7. Unresolved Issues

Thus far, we've seen that several worries that have been expressed about treatments of disagreement based on **Independence**-style principles may be assuaged. That is, of course a long way from offering a detailed Conciliationist recipe for accommodating the evidence provided by the disagreement of others. In this section, I'd like to note two of the issues facing this Conciliationist project.

a. Formulating a principle for belief-revision

As noted above, **Independence** needs to be supplemented by some principle describing how the dispute-independent assessment of the other person's epistemic credentials should inform one's response to disagreement. And principle (B), as stated, is still nothing like a general principle of this sort. For one thing, it ignores the question that we saw above to be important: whether the other person's expressed disagreement is sincere. More importantly, it's restricted to evidence that other person is one's epistemic *equal*, and this is obviously just a special case. Intuitively, to the extent that the dispute-neutral assessment gives reason to believe that the other person is one's superior, one should move one's belief farther toward that of the other. And even if the dispute-neutral assessment gives good reason for thinking the other person only slightly less qualified epistemically, some change of belief in the direction of her belief will often be warranted. Moreover, (B) says nothing about how much revision is required in any case; as it stands, it's not even clearly very Conciliatory. We might come closer to a general Conciliatory principle (at least for the most-discussed cases in which the other person expresses an opinion sharply opposed to one's own) as follows:

Revision In cases where one has strong confidence that P, and another person expresses equally strong confidence that $\sim P$, one should revise one's belief in the direction of the other person's expressed belief to the extent that

(A) One's dispute-independent assessment of her epistemic credentials yields high estimates for:

1. the likelihood that her expressed disagreement is sincere,
2. her degree of informedness, and
3. the likelihood of her having reasoned correctly from the evidence she has;

and

(B) The reasons for one's assessments of 1-3 are strong.

In cases where one's dispute-independent assessment yields very strong evidence that the other person is sincere, about equally well-informed, and about equally likely to have reasoned correctly from her evidence, one should cease to be much more confident that P than that $\sim P$.

Even here, however, many questions are left open: How far, in general, should one revise one's belief? How does the principle extend to cases where initial confidence is distributed differently—e.g. where both parties are more confident that P than $\sim P$, but to different degrees? As we saw above [fn.3], the initially attractive idea of uniformly splitting the difference in credences does not sit well with the motivations for Conciliationism.

So we are a long way from having a formula, or even a recipe with quantities, describing in general how one should react to disagreement. Given the complexity of the subject, and the early state of the discussion of this issue, this should not be surprising. And of course, Conciliationism is in no way special here; non-conciliationist writers have been equally unable to come up with precise general directions for accommodating disagreement evidence. But the burden of the arguments above was not to provide such directions. It has been to show that there is room for a Conciliationist approach to disagreement; in particular, to show that reliance on a

dispute-independent assessment of the other's belief's epistemic credentials does not have the untoward consequences it might seem to have.

b. Defining the scope of Independence²²

At the heart of Conciliationism, I've claimed, is the insight that in responding to the evidence of another's disagreement, one must avoid dismissing the other person's dissent in a question-begging way. I've formulated this as **Independence**, the requirement that I evaluate the epistemic credentials of the dissenting belief in a way that doesn't rely on the *reasoning* behind my initial belief that P.

A simpler proposal for preventing question-begging would have required independence only from P itself. But that would leave open the possibility of dismissing one's friend's beliefs in ways that would still intuitively beg the question. For example, suppose that in **Mental Math** my thinking went through five steps (dividing the bill by two, taking ten percent of that figure, multiplying that by two, adding the result of this to the result of step 1, and rounding up to the next dollar). I should not be able to dismiss my friend's belief by this sort of reasoning: "Our shares before rounding are \$42.87. So \$45 cannot be correct." Yet such reasoning would not depend on the belief my friend has disagreed with.

Unfortunately, this more inclusive formulation raises a question not adequately answered yet: exactly what sort of reasoning is part of "my reasoning behind P"? After all, my beliefs may be supported indirectly by many factors, especially in complex cases. Holists will insist that huge portions of my corpus of belief are relevant evidentially—they help support or rationalize my belief. Even the fact that I disbelieve certain claims that would defeat my justification for P is relevant. So do I have to put everything aside in assessing the credentials of my friend's dissent? Surely not—that would leave me unable to make any assessment of my friend's credentials.

One possible approach is suggested by the following variation of the example above: Suppose my friend agrees with all of the steps of my reasoning up to and including that the pre-rounding amount is \$42.87. Nevertheless, she insists that the answer is \$45, not \$43. In this case, it does not seem wrong for me to use the steps we agree about in evaluating her belief. In fact, the case then seems much like the extremely high rational confidence cases considered

²² The importance of this problem was emphasized to me by Jennifer Lackey.

above. It's not that I have extremely high rational confidence in P, but I do have extremely high rational confidence in the conditional of the last two steps in my reasoning, which my friend denies. Thus I have excellent reason to think that something screwy is going on. As before, personal information will strongly favor the hypothesis that the screwiness affects my friend rather than me. So the correct result will be achieved, and in an intuitively attractive way. This suggests that we understand the extent of the reasoning to be bracketed to be just that which is in dispute.

This is, I should emphasize, a very rough pass. It doesn't distinguish between steps which are denied by friend, and ones which simply aren't affirmed. It also leaves open questions about how to treat claims about which I don't know my friend's attitude: if we haven't discussed the steps of reasoning, I don't know for sure if she gets off the boat at the last step, or at some earlier step. There are some natural approaches for filling these gaps, but for now, I want only to acknowledge the complexities that lie ahead.

Again, it is important to note that the difficulty just described does not necessarily count against Conciliationism. That will depend on whether non-Conciliationist views can avoid encountering it, or whether they will also need to address the same problem. I'll turn to that question in the next section.

8. Beyond Conciliationism

The majority of this paper has been concerned to argue for the defensibility of Conciliatory approaches to the disagreement problem. In doing so, it has exposed some problems that a fully spelled-out Conciliationist account of disagreement would have to face. In this section, I want to argue that these problems--with defining the scope of **Independence**, and with incorporating the independent assessment into belief-revision--will have to be faced by any reasonable view. The problems, then, should be seen as indicating the difficulty of providing a general account of the epistemology of disagreement, rather than as posing a special obstacle to Conciliationism. Moreover, I'll argue that the issues raised here arise not only in the context of accounting for the epistemic significance of disagreement, but for a wide, but understudied, range of cases: those involving evidence that I may have reasoned improperly.

Of course, for radically Steadfast views, on which even the disagreement of my friend in the original **Mental Math** case gives no reason at all for modifying my confidence in \$43, there

is no need for worrying about what must be put aside in assessing my friend's credentials, or with how to take account of the resulting assessment. But such extreme versions of Steadfastness are highly implausible, and are not to my knowledge defended by anyone.²³

So suppose we consider a moderately Steadfast view: that disagreement, at least in cases like **Mental Math**, should make me significantly less confident of my original belief, but that in many ordinary cases of disagreement (say, in philosophy), my original reasoning (perhaps insofar as it was in fact correct) justifies me in thinking that my friend is substantially more likely to be mistaken than I am, and so not much revision is required. Can we obtain the right results on this sort of view, without relying at all on the sort of bracketing or putting aside of my original reasoning mandated by **Independence**?

The alternative would seem to be that, instead of *bracketing* my original reasoning, we would instead simply *combine* the reasons for P that support my initial belief together with the reasons for lowering confidence in P provided by the disagreement of my otherwise well-qualified friend. On this kind of view, my original reasons for belief in P are not put aside at all in determining what I should believe. To the extent that belief-change is needed, it's because those original reasons are partly undermined or outweighed by the reasons for doubting P provided by the disagreement.

It seems to me that this sort of view cannot quite be correct. To see why, consider again the **Mental Math** case. Suppose my friend and I are fully in agreement that our total bill is \$71.20, so the only question is what our shares of this amount are, with 20% tip and rounded up to the nearest dollar. And suppose that I've in fact arrived at the correct answer, \$43, by faultless appreciation of the relevant mathematical reasons. In this case, there's a sense in which the reasons for my initial belief are completely decisive: given the parameters of the problem, there's simply no possibility that \$45 could be right, and no possibility that \$43 could be wrong. Mathematical reasons are not like inductive reasons that can be weighed against or undermined by additional evidence. So insofar as I'm weighing my original reasons in "at full strength" (i.e., not putting them aside, at least to some extent, in my cognitive deliberations), and insofar as I combine all these reasons together correctly, it turns out that the evidence of my friend's

²³ Complete Steadfastness would hold that (at least in cases where my original calculations happened to be correct) I'd be rational in ignoring even hordes of friends who all were proven calculating whizzes, and all agreed that I had erred.

disagreement should in no way diminish my confidence in \$43. But this is not the result the moderately Steadfast theorist is after.²⁴

One might worry that something in this argument must be wrong. After all, even before learning of my friend's disagreement in **Mental Math**, I shouldn't have been absolutely confident in my answer. How, then, could those reasons be as decisive as the argument suggests?

The answer to this worry is that I *did* have reason to put my mathematical reasoning partially aside, even before learning of my friend's disagreement. After all, a major reason for being initially less-than-absolutely confident in \$43 is that I know that I and other people fairly commonly make mistakes in doing mental math. But again, if I were simply to combine the reasons for caution provided by this fact with the full-strength mathematical reasons, the fact about human fallibility would have no effect. Nothing about human fallibility affects the connection between the total bill's being \$71.20 and my share (with 20% tip, rounded up) being \$43. If I didn't at least partially put the mathematical reasons aside, then no amount of additional evidence would have any effect.²⁵

²⁴ One might question whether my initial reasons in this case are really decisive. If I represent those reasons purely subjectively—e.g., as “I, a generally reliable calculator, seem to remember getting \$43,” or “I now seem to see that \$43 is the correct answer,” then of course countervailing evidence can gain some purchase. But insofar as one acknowledges that my initial belief is made rational in part by the actual mathematical reasons which I correctly appreciate, one must acknowledge that my belief is supported decisively. The relation between the “data”—the amount of the bill and other parameters of the problem—and the claim that our shares are \$43 is not a relation that can be outweighed or undermined by anything.

²⁵ The contrast between reasoning in one's head and careful calculation has been cited by anti-Conciliationists as support for their view. As we've seen, for example, the effect of disagreement on mental-math-based beliefs is greater than the effect of disagreement on beliefs based on careful paper calculations. The anti-Conciliationist's suggestion is that the difference is to be explained in terms of the differences in the strength of the original reasoning, and thus that the original reasoning can't really have been put aside, contra **Independence**. But the considerations reviewed above suggest that the examples really pull in exactly the opposite direction. For the steps of reasoning one does in one's head may be exactly the same as the steps one does on paper. Insofar as those steps count as reasons for one's conclusion, they are equally (and maximally) strong. What differs between the cases is the second-order reasons for doubting the first-order reasoning. And those second-order reasons can only have purchase to the extent that one puts the first-order reasons aside.

This makes it clear that the general point about the necessity of (at least partially) bracketing some of our reasoning applies in wide areas outside the context of disagreement. There are countless possible reasons for me to be unsure of whether I've made a cognitive error, and thus unsure of conclusions I've reached. They range from knowledge of past errors that I or others have made, to unexplained feelings of dizziness, to memories of my past episodes of psychosis, or recent meals of magic mushrooms. The sort of evidence provided by disagreement of others is not really an exotic or peculiar epistemic phenomenon. It just involves another kind of indication one can have that one may have made a cognitive mistake.

It should also be noted that the necessity of putting aside one's original reasoning is not limited to cases in which that reasoning is deductive. The example of my having doubts about arithmetic calculation makes the point particularly clearly, taking advantage of the conclusive nature of deductive reasoning. But once the example is understood, its lessons clearly apply more generally. For what lies behind our verdict in the calculation case is our seeing that failure to put aside the reasoning under consideration amounts to begging the question of whether that reasoning was correct. And that is something we must not do, if we are to take seriously the possibility of our own cognitive error.

The bottom line, then, is this: Rationality requires that I take seriously evidence of my own possible cognitive malfunction in arriving at my beliefs. But insofar as I'm willing to do this, I must evaluate evidence for that possibility in a way that is (at least somewhat) independent of some of my reasoning. And I must then use this evaluation in arriving at my final belief. So any reasonable epistemology must face the problems we examined above: accounting for which portion of my original reasoning I must (at least somewhat) put aside in assessing the likelihood of my error, and determining how to integrate this assessment into my final beliefs. The two problems we saw facing the Conciliationist in evaluating and responding to disagreement-based evidence, then, both are highly general, in two different ways. For one thing, they are not just problems for Conciliationism: they are problems for every sensible account of the epistemology of disagreement. For another, they apply in areas quite removed from the disagreement issue.

The difficulty of these problems, I think, flows in part from a structural oddity that's inevitably involved in many sorts of epistemic self-criticism. On the one hand, one must evaluate one's beliefs in a way that does not simply beg the question of their cogency. On the other hand, one must do so from within the framework of one's beliefs—for one has no other

way to think about anything.²⁶ There is an obvious tension between these requirements, and it's far from clear that there will be a nice, clean recipe for resolving that tension in any arena. But resolving the tension somehow is an inescapable part of reacting rationally to the possibility of our own rational failures, something we're required to do all the time. Understanding how the tension is properly resolved, then, is essential to understanding how to cope rationally with a central aspect of our epistemic predicament.

References

- Christensen, D. (2007) "Epistemology of Disagreement: The Good News," *Philosophical Review* 116: 187-217.
- Elga, A. (2007), "Reflection and Disagreement," *Noûs* 41: 478-502.
- . (forthcoming), "How to Disagree about how to Disagree," in Feldman and Warfield,.
- Evnine, S. J. (2008), *Epistemic Dimensions of Personhood* (New York: Oxford University Press).
- Feldman, R. (2006), "Epistemological Puzzles about Disagreement," in S. Hetherington (ed.) *Epistemology Futures* (New York: Oxford University Press).
- (2007) "Reasonable Religious Disagreements," in L. Antony (ed.) *Philosophers without Gods: Meditations on Atheism and the Secular Life* (New York: Oxford University Press).
- Feldman, R. and T. Warfield, eds. (forthcoming), *Disagreement* (New York: Oxford University Press).
- Frances, B. (forthcoming), "The Reflective Epistemic Renegade," *Philosophy and Phenomenological Research*.
- Fumerton, R. (forthcoming) , "You Can't Trust a Philosopher," in Feldman and Warfield.
- Harman, G. (1986), *Change in View* (Cambridge, MA: MIT Press).

²⁶ See Evnine (2008), ch. 6, for interesting discussion of this general issue.

- Jehle, D. and B. Fitelson (2007), "What is the 'Equal Weight View'?", online at:
<<http://fitelson.org/ew.pdf>>
- Kelly, T. (2005) "The Epistemic Significance of Disagreement," *Oxford Studies in Epistemology* 1: 167 - 96.
- . (forthcoming), "Peer Disagreement and Higher-Order Evidence," in Feldman and Warfield.
- Kornblith, H. (forthcoming), "Belief in the Face of Controversy," in Feldman and Warfield.
- Lackey, J. (forthcoming a), "A Justificationist View of Disagreement's Epistemic Significance," in A. Haddock, A. Millar and D. Pritchard, eds., *Social Epistemology* (Oxford: Oxford University Press).
- . (forthcoming b), "What Should We Do when We Disagree?" in T. S. Gendler and J. Hawthorne, eds., *Oxford Studies in Epistemology III* (Oxford: Oxford University Press).
- Pettit, P. (2006), "When to Defer to Majority Testimony—and When Not," *Analysis* 66: 179-87.
- Shogenji, T. (2007) http://socrates.berkeley.edu/~fitelson/few/few_07/shogenji.pdf
- Sosa, E. (forthcoming), "The Epistemology of Disagreement," in his *Armchair Philosophy* (Princeton University Press, 2010).
- van Inwagen, P. (1996) "It is Wrong, Everywhere, Always, and for Anyone, to Believe Anything upon Insufficient Evidence," in *Philosophy of Religion: the Big Questions*, edited by Eleonore Stump and Michael J. Murray. New York: Blackwell (1999).