Reflection and Disagreement

Adam Elga*

Version of February 6, 2006

Introduction

There are well-informed experts and gurus: people to whom we should defer entirely. There are fakes and fools, who should be ignored. In between, there are friends and other advisors (including our own future and past selves), whose opinions should guide us in a less than fully authoritative way.

How, exactly, should we be guided by outside opinions?

Experts and gurus

Start with the simplest case: complete deference. When it comes to the weather, I completely defer to the opinions of my local weather forecaster. My probability for rain, given that her probability for rain is 60%, is also 60%. And the corresponding constraint holds for other propositions about the weather. Using a variant of Gaifman's (1988) terminology: I treat her as an **expert** about the weather. That means: conditional on her having probability x in any weather-proposition, my probability in that proposition is also x.[1]

In treating my forecaster this way, I defer to her in two respects. First, I defer to her *information*: "As far as the weather goes," I think to myself, "she's got all the information that I have—and more." Second, I defer to her *judgment*: I defer to the manner in which she forms opinions on the basis of her information.

[1]See also Hájek (2003, 311), to which I owe both the above example and also the notion of subject-matter-restricted expertise. Rich Thomason has pointed out that this way of restricting expertise to a subject matter can be at most a rough approximation. For there is no limit to the sort of evidence that might be relevant to the evaluation of weather-propositions. As a result, to treat the forecaster as an expert in the above sense is to potentially defer to her on the evaluation of any sort of evidence.

In the above case, we may suppose, I am right to treat my forecaster as an expert. But advisors don't always deserve such respect. For example, suppose that the forecaster has plenty of meteorological information, but I can see that she is dead drunk and so isn't responding properly to that information. In that case, I shouldn't treat her as an expert. Or suppose that the forecaster responds perfectly well to her information, but I can see that I have information that she lacks. In that case too, I shouldn't treat her as an expert.

Even in such cases, I shouldn't just ignore her opinion. How should I incorporate it? If my forecaster is drunk or otherwise addled, then I should only partially defer to her judgment. I postpone discussion of such cases. For now, suppose that I do completely defer to my forecaster's judgment. Nevertheless, I think that she lacks relevant information that I possess. What then?

An example will suggest the answer. Suppose that my forecaster lacks one highly relevant tidbit: that I have been secretly seeding the clouds for rain. Suppose that I'm sure her probability for rain is low—5%, say. In this case, I shouldn't set my probability for rain to that same low value, since my cloud-seeding activities make rain much more likely. But I *should* be guided by the forecaster's opinions. Roughly: my probability for rain should be what hers would be *if she were informed that I'd been seeding the clouds*.

More precisely: when I have information that my forecaster lacks, I shouldn't defer to her unconditional opinions. For those opinions are based on an impoverished evidential base. But I *should* defer to her *conditional* opinions: her opinions conditional on all of my extra information. When an agent defers to an advisor in this way, let us say that the agent treats the advisor as a **guru**.

Formally: suppose that I have probability function P. Then I treat an advisor as an **expert** if for any proposition H and for any probability function P' that I think the advisor may have,[2]

$$P(H|\text{advisor has } P') = P'(H).$$

---

[2]This is sometimes written in the (equivalent) form: $P(H|\text{advisor's prob for } H \text{ is } x) = x$.

In contrast, I treat the advisor as a **guru** if

$$P(H|\text{advisor has } P') = P'(H|X),$$

where X is my extra information, supposing that the advisor has probability function P'.[3,4]

Reflection

The above discussion concerns the manner in which one ought to take into account the opinions of others. But one should also take into account the opinions of one's own future self. How? A provocative and precise answer is given by the **Reflection Principle**[5], according to which one should treat one's future self as an expert. The Reflection Principle entails, for example, that one's current probability for rain, given that one will tomorrow have probability 25% for rain, ought also to be 25%. And it entails that the corresponding constraint should hold for all other propositions and time intervals.

But recall the weather forecaster case. To treat my forecaster as an expert is to defer to her with respect to both judgment and information. I shouldn't defer to her in this way if I have reason to doubt her on either score. And the same goes for my future self: if I expect that I won't be thinking straight next week, or if I expect that I will lose information between now and next week, then I *shouldn't* treat my next-week self as an expert (Christensen 1991, Skyrms 1987, Talbott 1991).

Reflection has been criticized on exactly these grounds. But we should distinguish two complaints. On the one hand, there is the complaint: rationality does not require that one defer to one's future *judgment*, since one may expect that one's judgment will become impaired. For example, Bob may reasonably doubt that at

---

[3]Compare to the motivation given in Hall (1994) for the move from the Old Principal Principle to the New Principal Principle.

[4]The above definition is appropriate only if countably many potential credence functions are in play. In the more general case, a fancier reformulation is in order. Such a reformulation might invoke a conception of conditional probability that allows for probabilities conditional on probability-zero propositions (Popper 1952, Renyi 1955). Or it might invoke integrals over probability densities. I suppress such elaborations here and in subsequent discussion.

[5]See van Fraassen (1984), van Fraassen (1995), Goldstein (1983).

midnight he will be in good shape to drive, even if he expects that by midnight he'll be drunkenly confident that his driving skills are perfect (Christensen 1991). In this case, Bob is reasonable, but violates Reflection. Against this complaint, I have no new help to offer the friends of Reflection.[6]

But there is a second complaint, to which there is an easy and satisfying response. The second complaint: it isn't a rational requirement that one defer to one's future self with respect to *information*, since one may expect to lose information. For example, Joe may reasonably be confident that this morning he had sushi for breakfast, even if he expects that in a year, he will have forgotten the truth of that claim (Talbott 1991). In this case, Joe is reasonable, but violates Reflection.

In response to this second complaint, it has been suggested that Reflection be limited so that it simply does not apply to cases of expected information loss.[7] Such a limited principle would avoid trouble by remaining silent about the trouble cases. But Reflection should not remain silent about such cases, since one's future opinion *should* constrain one's present opinion, even if one expects to lose information.

For a better response, think back to the previous section. Suppose that I defer to the judgment of my weather forecaster. Then, even if I have information that she lacks, her opinion *should* constrain my own. Roughly, I should believe what I think she would believe, if she were given my extra information. More precisely: I should treat my forecaster not as an expert, but a guru.

The same holds with respect to my future self. Suppose that I defer to the judgment of my future self. Then, even if I have information that I will later lack, my future opinion *should* constrain my present opinion. Roughly, I should believe what I think my future self would believe, if my future self were given my present extra information. More precisely: I should treat my future self not as an expert, but as a guru.

The moral is that expected information loss does not *break* the connection between one's future opinion and one's present opinion: it *modifies* that connection. The modified connection is expressed by the principle "Treat your future self as a

---

[6]Some lines of response are explored in van Fraassen (1995).

[7]Such limited principles include "Reflection Restricted", from Jeffrey (1988, 233), and "Confidence", from Hall (1999, 668).

guru".[8]

This modified principle handles cases of information loss not by remaining silent, but by yielding reasonable verdicts. For example, recall the case of Joe, the forgetful sushi-eater. The modified principle yields the desired verdict, that Joe should be confident that he had sushi for breakfast. For though Joe expects his future self to doubt this claim, he expects his future self to be confident in it, *conditional on his current extra information* (which includes vivid memories of having recently eaten sushi).

There is another sort of information loss, a sort associated with losing track of who one is or what time it is. Information loss of that sort can also lead to violations of Reflection. For example[9], suppose that you are waiting for a train. You are only 50% confident that the train will ever arrive, but are sure that if it does arrive, it will arrive in exactly one hour. Since you have no watch, when fifty-five minutes have in fact elapsed you will be unsure whether an hour has elapsed. As a result, at that time you will have reduced confidence—say, only 40% confidence—that the train will ever arrive. So at the start, you can be sure that when fifty-five minutes have elapsed, your probability that the train will ever arrive will have gone down to 40%. But that doesn't show that your current probability should be 40%. So your anticipated imperfect ability to keep track of time creates a violation of Reflection.[10]

Again, the proper lesson is not that the connection between current and future beliefs is broken, but rather that it is modified. For example, in the above case your probabilities should not match what you expect to believe in fifty-five minutes. Instead they should match what you expect to believe in fifty-five minutes *given that exactly fifty-five minutes have elapsed*. More generally, the definition of what it is to treat someone as a guru can be modified in order to "bracket off" the manner

---

[8]Formally, the proposal is that for any proposition H and for any probability function P' that one thinks one may have at a future time t,

$$P(H|\text{I have } P' \text{ at } t) = P'(H|X),$$

where X is the extra information one presently possesses, supposing that one has P' at t. For a similar proposal, along with criticisms of the present one, see Weisberg (2005, Section 5).

[9]This example is a retelling of the "prisoner in the cell" case from Arntzenius (2003).

[10]Another example: Horgan (2004) convincingly diagnoses the Reflection-violating belief change associated with the Sleeping Beauty problem as resulting from information loss about what time it is.

in which an agent's uncertainty about what time it is (and who she is) affects her beliefs about other matters.[11]

So: when one completely defers to one's future judgment, one should treat one's future self as a guru. But when should one completely defer to one's future judgment? More generally, when should one completely defer to the judgment of any advisor?

Experts and gurus are rare

Rarely.

Only in highly idealized circumstances is it reasonable to defer to someone's opinion *absolutely whatever* that opinion might be. For example, upon finding out that my forecaster is confident that it will snow tomorrow, I will follow suit. But upon finding out that my forecaster is confident that it will rain eggplants tomorrow, I will not follow suit. I will conclude that my forecaster is crazy. The same goes for the news that I myself will believe that it will rain eggplants tomorrow. In realistic cases, one reasonably discounts opinions that fall outside an appropriate range.

In addition, not even a perfect advisor deserves absolute trust, since one should be less than certain of one's own ability to *identify* good advisors.[12]

So: only in highly idealized cases is it appropriate to treat someone as an expert or a guru, and so to completely defer to that person's judgment. All the more reason to consider cases of partial deference, to which we now turn.

Dividing the question

How should one take into account the opinions of an advisor who may have imperfect judgment? That question factors into two parts:

---

[11]Applied to the case of a single person over time, the resulting principle requires that for any subject S with probability function P at time t, any proposition H, and any probability function P' that the subject thinks she might have at future time t',

$$P(H|\text{I have } P' \text{ at } t \text{ \& I am } S \text{ at } t) = P'(H|X \text{\& I am } S \text{ at } t'),$$

where X is the subject's extra information at time t, on the supposition that she has P' at t'. See also Schervish et al. (2004).

[12]I owe this point to Peter Railton.

1. To what degree should one defer to a given advisor's judgment? For example, when should one count an advisor's judgment as completely worthless? Or as approximately as good as one's own? Or as better than one's own, but still less than perfect?

2. Given one's assessment of an advisor's level of competence, how should one take that advisor's opinion into account?

On the first question, I have no substantive answer to offer here. My excuse is that the question concerns a huge, difficult, and domain-specific matter. How should one judge the epistemic abilities of weather forecasters, dentists, math professors, gossipy neighbors, and so on? This is a question with the same sort of massive scope as the question: "When does a batch of evidence support a given hypothesis?" Fearsome questions both, and worthy of investigation. But leave them for another day.

Here I will focus on the second question. Assume that you defer to an advisor's judgment to a certain degree. *Given* that rating of the advisor's judgment, how should you take her opinions into account? We have already settled this in the special case in which you utterly defer to the advisor's judgment. In that case, you should treat the advisor as a guru. It remains to consider cases in which you defer to an advisor's judgment only partially.

Start with the simplest such case, a case that has been the subject of considerable discussion and dispute (Christensen 2004, Feldman 2004, Kelly 2005, Plantinga 2000, van Inwagen 1996): how should your opinion be guided by an advisor who you count as having judgment that is as good as your own?

Peer disagreement: setup

Suppose that you and your friend independently evaluate the same factual claim—for example, the claim that the death penalty significantly deterred crime in Texas in the 1980s.[13] Each of you has access to the same crime statistics, sociological reports, and so on, and has no other relevant evidence. Furthermore, you count your

---

[13]I borrow this example from van Inwagen (1996, 141).

friend as an **epistemic peer**—as being as good as you at evaluating such claims.[14,15]

You perform your evaluation, and come to a conclusion about the claim. But then you find out that your friend has come to the opposite conclusion. How much should this news move you in the direction of her view? Should you always give your friend's assessment equal weight, and think that it is no more likely that you're right than that she is? Or can it sometimes be rational for you to stick to your guns, or at least give your own assessment *some* extra weight?

Answer: you should give the assessments equal weight.

Before refining and defending this **equal weight view**, let me attack it.

Unwelcome consequences of the equal weight view

According to the equal weight view, one should give the same weight to one's own assessments as one gives to the assessments of those one counts as one's epistemic peers. If the view is right for the case of one peer, it surely is also right for the case of many peers.[16] But in the case of many peers, the view seems to have unwelcome consequences.

First unwelcome consequence: spinelessness. Consider an issue on which you count many of your associates as epistemic peers. If the issue is at all tricky, your peers undoubtedly take a wide spectrum of stances on it. (This is especially true if your peers are philosophers.) The equal weight view then requires you to weigh each stance equally, along with your own. But that requires you to think, of each stance, that it is very unlikely to be right. Typically, it will follow that you ought to suspend judgment on the issue. Since it seems that you are in this circumstance with respect to a great many issues, the equal weight view requires you to suspend

---

[14]I owe the term "epistemic peer" to Gutting (1982) by way of Kelly (2005), though I use it in a different way than they do. See note 22.

[15]Note that in setting up the problem, the initial assumption is that you *count* your friend as your epistemic peer. That contrasts with some presentations, in which the initial assumption is that your friend *is* your epistemic peer. The former assumption is appropriate, however. For example, one sometimes is reasonable in thinking wise advisors to be foolish. Evidence, after all, can be misleading. In such cases, one is reasonable in being guided by one's *assessments* of the advisor's ability, even if those assessments are in fact incorrect.

[16]Or at least: a very natural generalization of it is right for the case of many peers. The generalization would have to take into account, for example, the degree to which one judges that one's peers reached their conclusions independently. For simplicity, I suppress such complications here.

judgment on all of these. Do you have any convictions on controversial political, philosophical, or scientific matters? The equal weight view seems to say: kiss them goodbye. It is implausible that rationality requires such spinelessness (Pettit 2005, van Inwagen 1996).

Second unwelcome consequence: lack of self-trust. Suppose that a great fraction of those you count as your epistemic peers agree with you on some issue. Then the equal weight view says: stick with your initial assessment. Great! Except that the *reason* for sticking to that assessment has very little to do with your own evaluation of the common stock of evidence, and very much to do with the fraction of your peers who agree with you. Shouldn't your own careful consideration of the issue count for more than 1/50th, even if you have 50 epistemic peers? If not, then one might just as well form views on controversial matters by simply sending out a poll (Landesman 2000). It is implausible that rationality requires you to give your own consideration of the issue such a minor role.[17]

These are unwelcome consequences of the equal weight view. One might try to mitigate the consequences by claiming that most of the disputed issues are not matters of fact, or that people rarely share their relevant evidence, or that one shouldn't count many of one's associates as peers after all.[18] Or one might try to make the consequences more palatable (Christensen 2004, Feldman 2004). I will not discuss these strategies at present. The present point is this: even if the equal weight view is right, it is not *obviously* right.

What reinforces this point is that the equal weight view stands in competition with views that clearly avoid the unwelcome consequences. Two such views deserve special note.

Competitors to the equal weight view

The first competing view is the **extra weight view**, according to which one should give one's own assessment more weight than the assessments of those one counts

---

[17]Furthermore, in light of such considerations, one might be tempted to avoid the hard work of thinking an issue through oneself, by simply deferring to the aggregated opinions of one's peers. This leads to the free-rider problem explored in List and Pettit (2003).

[18]van Inwagen (1996) scotches many such strategies.

as epistemic peers.[19] For example, when you find out that your sole epistemic peer has arrived at a contrary conclusion, the extra weight view says that you should be pulled a *bit* in the peer's direction—but not half way. You should still think it more likely that you are right than that the peer is.

By granting special status to one's own assessments, the extra weight view mitigates the unwelcome consequences described in the previous section. For if your own assessment gets extra weight, you may reasonably stick to your guns to a great degree, even if a handful of your peers disagree. So the extra weight view does not require a spineless suspension of judgment on all controversial matters. Furthermore, on the extra weight view, your own assessment of the evidence has more impact on the formation of your overall view than it does on the equal weight view.

To introduce another competitor to the equal weight view, recall the guiding question: when you find out that your peer has arrived at the opposite conclusion as you, how much should you be moved?

According to the **right-reasons view**, the answer depends on how good your initial evaluation of the evidence was.[20] For example, suppose that the shared batch of evidence *in fact* strongly supports the disputed claim. You correctly apprehend this, but your peer misjudges the force of the evidence and as a result disagrees. In this case, the right-reasons view says it can be reasonable for you to stick to your (correct) evaluation.

What motivates the view is that there is an asymmetry in the above sort of dispute. Although each disputant *thinks* that he has correctly assessed the force of the evidence, at most one of the disputants has *in fact* done so (Kelly 2005, 180). A defender of the right-reasons view invites us to think of things as follows. Suppose that you have arrived at the conclusion that the evidence in fact supports—

---

[19]This view is described (but not endorsed) in Feldman (2004, 14).

[20]I learned of this view from Kelly (2005, 180). Here is a representative quotation:

> The rationality of the parties engaged in [a disagreement among epistemic peers] will typically depend on who has *in fact* correctly evaluated the available evidence and who has not. If you and I have access to the same body of evidence but draw different conclusions, which one of us is being more reasonable (if either) will typically depend on which of the different conclusions (if either) is in fact better supported by that body of evidence.

call it Conclusion C. Your peer has gotten things wrong and arrived at some other conclusion. What does rationality require? That you *stick* with Conclusion C, and that your peer *switch* to Conclusion C.[21]

The right-reasons view mitigates the problems of spinelessness and lack of self-trust. For when you have in fact assessed the evidence correctly, the view entails that you may reasonably stick to that assessment, even if a handful of your peers disagree. For the same reason, the view allows that even if you have many peers, your own assessment of the evidence need not be swamped.

To sum up: when you disagree with your peer, the situation is symmetric in the sense that your assessment seems right to you, and your peer's assessment seems right to him. What might break the symmetry? The extra weight view says: you should give your own assessment extra weight because it is *yours*. The right-reasons view says: if in fact your assessment is right, it deserves extra weight because it is *right*. Both views avoid the unwelcome consequence that your own assessments must be swamped when you have a large number of peers.

In contrast, the equal weight view says: even though your assessment is your own, and even if it is in fact correct, you shouldn't favor it even a tiny bit. It is time to defend the equal weight view.

For the equal weight view: bootstrapping

To see the correctness of the equal weight view, start with a case of perceptual disagreement.

You and a friend are to judge the same contest, a race between Horse A and Horse B. Initially, you think that your friend is as good as you at judging such races. In other words, you think that in case of disagreement about the race, the two of you are equally likely to be mistaken. The race is run, and the two of you form independent judgments. As it happens, you become confident that Horse A won, and your friend becomes equally confident that Horse B won.

---

[21]In its most extreme form, the right-reasons view holds that when a disputant has correctly assessed the force of the evidence, rationality permits him to be *entirely unmoved* by the news of peer disagreement. A more modest version holds that when a disputant has correctly assessed the force of the evidence, rationality permits him to be moved *less than half way* in the direction of a peer's contrary assessment.

When you learn of your friend's opposing judgment, you should think that the two of you are equally likely to be correct. For suppose not—suppose it were reasonable for you to be, say, 70% confident that you are correct. Then you would have gotten some evidence that you are a better judge than your friend, since you would have gotten some evidence that you judged this race correctly, while she misjudged it. But that is absurd. It is absurd that in this situation you get any evidence that you are a better judge.

To make this absurdity more apparent, suppose that you and your friend independently judge the same long series of races. You are then allowed to compare your friend's judgments to your own. (You are given no outside information about the race outcomes.) Suppose for *reductio* that in each case of disagreement, you should be 70% confident that you are correct. It follows that over the course of many disagreements, you should end up extremely confident that you have a better track record than your friend. As a result, you should end up extremely confident that you are a better judge. But that is absurd. Without some antecedent reason to think that you are a better judge, the disagreements between you and your friend are no evidence that she has made most of the mistakes.

Furthermore, the above judgment of absurdity is independent of who *in fact* has done a better job. Even if in fact you have judged the series of races much more accurately than your friend, simply comparing judgments with your friend gives you no evidence that you have done so.

Here is the bottom line. When you find out that you and your friend have come to opposite conclusions about a race, you should think that the two of you are equally likely to be correct.

The same goes for other sorts of disagreements.

Suppose that instead of judging a race, you and your friend are to judge the truth of a claim, based on the same batch of evidence. Initially, you count your friend as an epistemic peer—you think that she is about as good as you at judging the claim. In other words, you think that in case of disagreement about the claim, the two of you are equally likely to be mistaken.[22] Then the two of you perform your

---

[22]My use of the term "epistemic peer" is nonstandard. On my usage, you count your friend as an epistemic peer with respect to an about-to-be-judged claim if only if you think that in case of disagreement about the claim, the two of you are equally likely to be mistaken. On more standard

12

evaluations. As it happens, you become confident that the claim is true, and your friend becomes equally confident that it is false.

When you learn of your friend's opposing judgment, you should think that the two of you are equally likely to be correct. The reason is the same as before. If it were reasonable for you to give your own evaluation extra weight—if it were reasonable to be more than 50% confident that you are right—then you would have gotten some evidence that you are a better evaluator than your friend. But that is absurd.

Again, the absurdity is made more apparent if we imagine that you and your friend evaluate the same long series of claims. Suppose for *reductio* that whenever the two of you disagree, you should be, say, 70% confident that your friend is the mistaken one. It follows that over the course of many disagreements, you should end up extremely confident that you have a better track record than your friend. As a result, you should end up extremely confident that you are a better evaluator. But that is absurd. Without some antecedent reason to think that you are a better evaluator, the disagreements between you and your friend are no evidence that she has made most of the mistakes.

Again, this absurdity is independent of who has in fact evaluated the claims properly. Even if in fact you have done a much better job than your friend at evaluating the claims, simply comparing your verdicts to those of your friend gives you no evidence that this is so.

The above argument has the same form as a certain well-known objection to reliabilism (Goldman 1986, Sosa 1997). According to reliabilism, one can gain knowledge by a reliable method, even if one does not know that the method is reliable. For example, suppose that your color vision is in fact reliable. In that case—according to reliabilism—by looking at a red wall you can come to know

usages, an epistemic peer is defined to be an equal with respect to such factors as "intelligence, perspicacity, honesty, thoroughness, and other relevant epistemic virtues" (Gutting 1982, 83), "familiarity with the evidence and arguments which bear on [the relevant] question", and "general epistemic virtues such as intelligence, thoughtfulness, and freedom from bias" (Kelly 2005). In defense of my use, suppose that you think that in case of disagreement about a claim, your friend is more likely than you to be mistaken. Then however intelligent, perspicacious, honest, thorough, well-informed, and unbiased you may think your friend is, it would seem odd to count her as an epistemic peer with respect to that claim, at least on that occasion. You think that in case of disagreement, she is more likely to get things wrong!

that the wall is red, even if you don't know that your vision is reliable.[23]

Now for the objection. If reliabilism is right, then in the above case you get some evidence that your vision is reliable merely by looking at the wall. That is because you come to know that your vision has operated correctly on that occasion. By looking at many walls, you can come to know that your vision has operated correctly on many occasions, and hence can come to know that your vision is reliable. In other words, you can come to know that your vision is reliable merely by checking that the outputs of your visual system agree with ... the outputs of your visual system. Such procedures, called "bootstrapping" (Vogel 2000, 615) or the gaining of "easy knowledge" (Cohen 2002), are clearly illegitimate. But it follows from reliabilism they can be legitimate. That is the objection.

We have seen that rivals to the equal weight view are subject to an analogous objection. For suppose that it was legitimate to give your own evaluations more weight than those of a friend who you initially count as a peer. Then it could be legitimate for you to "bootstrap"—to come to be confident that you are a better evaluator than the friend merely by noting cases of disagreement, and taking it that the friend made most of the errors.[24] But that is absurd. So it is not legitimate to give your own evaluations more weight than those who you count as peers. A similar argument shows that it is not legitimate to give your own evaluations less weight than those who you count as peers.

So the equal weight view is correct.

The equal weight view, refined

This section describes two complications. They require tweaking the equal weight view, but not in a way that interferes with the above defense.

To start, recall what the equal weight view says. Suppose that before evaluating a claim, you think that you and your friend are equally likely to evaluate it correctly. When you find out that your friend disagrees with your verdict, how likely should

---

[23]For ease of exposition, a simplified version of reliabilism about knowledge is targeted here. But the guiding idea of the objection can be wielded against more sophisticated versions.

[24]Proponents of the right-reasons view are only committed to the legitimacy of this sort of bootstrapping when the bootstrapper does in fact evaluate the evidence better than her opponent. The bootstrapping conclusion is absurd even in this case.

you think it that you are correct? The equal weight view says: 50%.

But here is a complication. Suppose that while evaluating the claim, you get some relevant information about the circumstances of the disagreement. For example, suppose that the weather gets extremely hot, and you know that your friend—unlike you—can't think straight in hot weather. In that case, when you find out that your friend disagrees, you should end up fairly confident that she is the mistaken one.

In particular, suppose that before evaluating the claim, the following was true: conditional on a disagreement arising in hot weather, you were 80% confident that your friend would be the mistaken one. In that case, when a disagreement *does* arise in hot weather, you should be 80% confident that your friend is the mistaken one (Christensen 2004, Section 4). And the same goes for initial degrees of confidence other than 80%.

The point is that you should not be guided by your prior assessment of your friend's *overall* judging ability. Rather, you should be guided by your prior assessment of her judging ability *conditional on what you later learn about the judging conditions*.

Notice that the above view is more general than the equal weight view, as first stated. For at the start, the equal weight view applied only to cases in which you initially count your advisor as a peer—as equally likely to be right, on the supposition that the two of you end up disagreeing. But the modified view also applies to cases in which you initially count an advisor as an epistemic superior—as being more than 50% likely to be right, on the supposition that the two of you end up disagreeing. Likewise, the view applies to cases in which you initially count an advisor as an epistemic inferior.

There is a second complication. The above view appeals to your prior assessment of your friend's abilities—the assessment you had before thinking through the disputed issue. But what if you thought through the disputed issue years ago, before you even met this friend? Then it won't help to consider what you believed about the friend's abilities way back then. For at that time, you had not even met the friend (and hence had no informed opinion of her abilities). So the equal weight view is useless in this case. How can it be fixed?

For an answer, notice that the whole point of considering your past beliefs was to get access to a state of opinion untainted by your detailed reasoning about the disputed issue.[25] *One* way to do that is to consider what you believed before thinking the issue through. But that is not the only way. Sometimes we may sensibly ask what a given agent believes, *bracketing* or *factoring off* or *setting aside* certain considerations. For example, suppose that your views on the trustworthiness of Jennifer Lopez derive from both tabloid reports and face-to-face interactions. In this case, we may sensibly ask what your views of Lopez are, setting aside what the tabloids say. To ask this is not to ask about your actual beliefs at some previous time. Rather, it is to ask what happens when we remove or extract tabloid-based information from your current state of belief.

Likewise, in case of disagreement between you and a friend, we may ask what you believe, setting aside your detailed reasoning (and what you know of your friend's reasoning) about the disputed issue. In particular, we may ask who you think would likely be correct, setting that reasoning aside. By construction, the resulting belief state is untainted by ("prior to") your reasoning about the disputed issue. But since *only* the disputed reasoning has been extracted, that belief state still reflects your general information about your friend's abilities. The equal weight view is best understood as invoking this non-temporal notion of prior belief. Here is the resulting view:

Equal weight view Upon finding out that an advisor disagrees, your probability that you are right should equal your prior conditional probability that you would be right. Prior to what? Prior to your thinking through the disputed issue, and finding out what the advisor thinks of it. Conditional on what? On

---

[25]Compare to the requirement of Christensen (2004) that explanations of a disagreement be evaluated independently from the disputed issue.

16

whatever you have learned about the circumstances of the disagreement.[26,27]

That completes my explanation of the equal weight view, an account of how to incorporate the opinions of an advisor to whom one only partially defers. But it does not complete my defense. For recall that the view seemed to have some very unwelcome consequences.

The problem of spinelessness

It is time to face the problem of spinelessness—the problem that an egalitarian view on how to respond to disagreement will recommend a spineless suspension of judgment on virtually all controversial issues. Let me pose the problem again, from a slightly different angle.

Views on disagreement face pressure from two directions. On the one hand, when one considers clean, pure examples of disagreement, it seems obvious that something like the equal weight view is correct. It seems obvious that when you

---

[26]Note that the circumstances of a disagreement should not include a detailed specification of the chain of reasoning that led you to your conclusion. For if they did, then making the relevant conditional probability judgment would involve thinking through the disputed issue—and hence would not be prior to your doing so.

For example, suppose that you and a friend get different answers to the same multiplication problem. In applying the equal weight view to this case, the circumstances of disagreement might include such factors as: the amount of scratch paper available to you and your friend; how much coffee each of you has recently drunk; how confident the two of you were in your respective answers (after doing the calculation, but before finding out about the disagreement); and whether the calculation involves carrying many 1s (this last factor would be relevant if, for example, you know your friend often forgets to carry his 1s).

In general, circumstances of disagreement should be individuated coarsely enough so that the relevant conditional probability judgment is genuinely prior to your reasoning about the disputed issue.

[27]In talking about your conditional probability that one is "right", the above formulation assumes that disputants arrive at all-or-nothing assessments of the claim under dispute. That assumption is adopted only as a convenience. It is relaxed in the following more general formulation of the view:

> Your probability in a given disputed claim should equal your prior conditional probability in that claim. Prior to what? Prior to your thinking through the claim, and finding out what your advisor thinks of it. Conditional on what? On whatever you have learned about the circumstances of how you and your advisor have evaluated the claim.

Note that this formulation governs one's response not just to cases in which your advisor disagrees, but also to cases in which she agrees or suspends judgment. Note also that this formulation does not presuppose any view on the "uniqueness thesis" (Feldman 2004), according to which rational disagreements can only arise from differences in evidence.

disagree about an arithmetic problem with a friend who you previously considered to be equally good at arithmetic, you should think yourself no more likely to be correct than your friend.

On the other hand, when one considers messy examples of real-world disagreements about hard issues, the equal weight view seems to lead to absurdity. Example: your friends take a range of stances on some basic political or ethical claim. By your lights, these friends are just as thoughtful, well-informed, quick-witted, and intellectually honest as you. Still, it seems obviously wrong that you are thereby required to suspend judgment on the claim, as the equal weight view seems to entail. To require this would be to require you to suspend of judgment on almost everything.

So: with respect to the clean, pure cases, there is pressure in the direction of the equal weight view. With respect to messy real-world cases, the equal weight view seems to lead to absurdity. What gives?

The answer is that the equal weight view does not lead to absurdity, because there is a relevant difference between the two sorts of cases. The difference is that in the clean cases one is in a position to count one's associates as peers *based on reasoning that is independent of the disputed issue*. But in the messy real-world cases, one is rarely in a position to do so. That is because in the messy cases, one's reasoning about the disputed issue is tangled up with one's reasoning about many other matters (Pettit 2005). As a result, in real-world cases one tends not to count one's dissenting associates—however smart and well-informed—as epistemic peers.

Let me explain, by way of a few examples.

Suppose that you disagree with your friend about the multiplication problem "What is 5243324×922?". You nevertheless count your friend as a peer: setting aside your reasoning about this particular problem, you think that she is equally likely to get the right answer in case of disagreement.[28] Your evaluation may be based on such factors as your friend's mathematical track record, what sort of training she has had, and so on. It need not based on any particular view on the value of

---

[28]Attempts to represent uncertainty about mathematics face the so-called "problem of logical omniscience" (Stalnaker 1991). For present purposes, it does no harm to treat arithmetical claims as logically independent contingent claims in the manner of Garber (1983). Arithmetical reasoning will simply be assumed to change these probabilities over time, in a way not represented in the model.

5243324×922, or on the answers to similar multiplication problems.[29]

In messy real-world cases, such independent peer evaluations are often unavailable. For example, consider Ann and Beth, two friends who stand at opposite ends of the political spectrum. Consider the claim that abortion is morally permissible.[30] Does Ann consider Beth a peer with respect to this claim? That is: setting aside her own reasoning about the abortion claim (and Beth's contrary view about it), does Ann think Beth would be just as likely as her to get things right?

The answer is "no". For (let us suppose) Ann and Beth have discussed claims closely linked to the abortion claim. They have discussed, for example, whether human beings have souls, whether it is permissible to withhold treatment from certain terminally ill infants, and whether rights figure prominently in a correct ethical theory. By Ann's lights, Beth has reached wrong conclusions about most of these closely related questions. As a result, even setting aside her own reasoning about the abortion claim, Ann thinks it unlikely that Beth would be right in case the two of them disagree about abortion.

In other words, setting aside Ann's reasoning about abortion does not set aside her reasoning about allied issues. And by Ann's lights, the accuracy of an advisor's views on these allied issues indicates how accurate the advisor is likely to be, when it comes to abortion. The upshot is that Ann does not consider Beth an epistemic peer with respect to the abortion claim.

So the abortion case is quite different than the multiplication case. Furthermore, the contrast between them is representative. In the clean, pure cases of disagreement used to motivate the equal weight view, the disputed issues are relatively separable from other controversial matters.[31] As a result, the agents in those examples often count their smart friends and associates as peers about the issues under dispute. But in messy real-world cases, the disputed issues are tangled in clusters of contro-

---

[29]The situation is different in case of disagreement about exceedingly simple mathematical or logical problems. When a friend disagrees about whether $1 + 1 = 2$, one may well not count her as a peer. For one's views on such a simple problem are closely linked to one's basic reasoning. For that reason, there may be no determinate fact about what one believes, setting aside one's views on whether $1 + 1 = 2$. For further discussion of this point, see the reply to the objection on p. 21.

[30]If you think that moral claims such as this are not factual claims, then please substitute a clearly factual, hotly contested political claim. Unfortunately, many such claims exist.

[31]Such pure cases include the savant cases from Feldman (2004) and Moller (2004), and the "split the check" case from Christensen (2004).

versy. As a result, though agents in those examples may count their associates as thoughtful, well-informed, quick-witted, and so on, they often do *not* count those associates as peers. For example, Ann does not count Beth as a peer with respect to the abortion claim.

Think of a smart and well-informed friend who has a basic political framework diametrically opposed to your own. Imagine that the two of you are both presented with an unfamiliar and tricky political claim. You haven't thought things through yet, and so have no idea what you will eventually decide about the claim. Still— don't you think that you are more likely than your friend to correctly judge the claim, supposing that the two of you end up disagreeing? If so, then however quick-witted, well-informed, intellectually honest, and thorough you think your friend is, you do not count her as an epistemic peer with respect to that claim. And if you do not count her as a peer, the equal weight view does not require you give her conclusion the same weight as your own. Indeed, if you think that your friend has been consistently enough mistaken about allied issues, then the equal weight view requires you to become *more* confident in your initial conclusion once you find out that she disagrees.

At the other extreme, think of a smart friend who has a basic political framework extremely similar to your own. Again, imagine that both of you have just been presented with an unfamiliar political claim. In this case, you may well think that in case of disagreement, your friend is just as likely as you to be correct. If so, and if you and your friend end up coming to opposite verdicts, then the equal weight view requires you to think it just as likely that she is right as that you are. But notice that friends like these—friends who agree with you on issues closely linked to the one in question—will very often agree with you on the one in question as well.

Moral: with respect to many controversial issues, the associates who one counts as peers tend to have views that are similar to one's own. That is why—contrary to initial impressions—the equal weight view does not require one to suspend judgment on everything controversial.

That is how the equal weight view escapes the problem of spinelessness.

Objections

Objection. The equal weight view escapes the problem of spinelessness only by advocating an ostrich-like policy of listening only to those with whom one already agrees.

Reply. First, everything said so far is compatible with the advice: listen to opposing arguments with an open mind. Everyone agrees that one should do that. At issue is the degree to which the mere fact of disagreement should change one's position. In other words, at issue is how much one should be moved when one either doesn't know the reasoning behind an advisor's dissenting conclusion, or does know the reasoning, but doesn't find it compelling.

Second, the equal weight view often requires one to be moved a fair bit by dissenting opinion. For though controversial matters tend to be linked together in clusters, it is not as though everyone walks in lockstep with their political, scientific, or philosophical clique. For example, imagine that you and a friend are presented with a tricky claim that you have not yet thought through. You may well think that your friend is almost as likely as you to evaluate the claim correctly, even if the two of you differ a bit on allied issues. If so, and if you and your friend end up disagreeing, then the equal weight view requires you to be significantly moved.

So the equal weight view doesn't say: bury your head in the sand. It does say: defer to an advisor in proportion to your prior probability that the advisor would be correct in case of disagreement. In practice, this means deferring most to advisors whose views (on matters closely linked to the issue in question) are similar to one's own.[32]

Objection. Return to the case of Ann and Beth, friends at opposite ends of the political spectrum. Ann counts Beth as less than an epistemic peer when it comes to abortion. She does so because she judges that Beth has gone wrong on many issues linked to abortion. But that judgment rests on the assumption that Ann's views on the linked issues are correct—an assumption to which Ann is not entitled.

Rather than taking her views on the surrounding issues for granted, Ann should

---

[32]Compare to Foley (2001, 105), which argues that one has (defeasible) reason to trust the beliefs of an anonymous advisor, on the grounds that there are "broad commonalities in the intellectual equipment and environment of peoples across times and cultures".

attend to the larger disagreement between her and Beth: disagreement about a whole cluster of issues linked to abortion. Ann should think of this whole cluster as a single compound issue, and should take into account Beth's disagreement about *that*. When she does so, she can no longer penalize Beth for going wrong on surrounding issues. So the equal weight view entails that Ann should suspend judgment about abortion, in light of Beth's disagreement. Furthermore, similar reasoning applies in many cases of real-world disagreement. So the equal weight view does after all require suspension of judgment on virtually everything controversial.

Reply. Consider the cluster of issues linked to abortion. Contrary to what the objection supposes, Ann does *not* consider Beth a peer about that cluster. In other words, setting aside her reasoning about the issues in the cluster, and setting aside Beth's opinions about those issues, Ann does not think Beth would be just as likely as her to get things right. That is because there is no fact of the matter about Ann's opinion of Beth, once so many of Ann's considerations have been set aside. Hence the equal weight view does not require Ann to suspend judgment about the cluster. That blocks the objection.

But why is there no fact of the matter about Ann's opinion, setting aside her reasoning about the cluster? To see why, return to the example in which you know Jennifer Lopez both from face-to-face interactions and tabloid reports. In that example, there is a determinate answer to the question: "What is your opinion of Lopez, setting aside what the tabloids say?" That is because there is a relatively self-contained path along which the tabloids influence your beliefs about Lopez. As a result, there is a natural way of factoring your belief state into a "prior" state, together with some additional tabloid-based information.

In contrast, there is no determinate answer to the question: "What is your opinion of Lopez, setting aside that humans have bodies and that the Earth exists?" That is because there is no natural way of factoring your belief state into a "prior" state, together with that additional information. Setting aside that humans have bodies and that the Earth exists, how confident are you that Lopez dated Ben Affleck? Or that one of Lopez's movies was recently panned by critics? Or that Hollywood even exists? These questions have no answers because the information to be set aside is enmeshed in too much of your reasoning to be cleanly factored off.

The same goes for Ann's reasoning about the cluster of issues linked to abortion. That cluster includes a wide range of issues: whether humans have souls, the age at which humans begin feeling pain, whether rights figure prominently in a correct ethical theory, and so on. To set aside Ann's reasoning about all of these issues is to set aside a large and central chunk of her ethical and political outlook. Once so much has been set aside, there is no determinate fact about what opinion of Beth remains. As a result, the equal weight view issues no determinate verdict about how Ann should respond to the larger disagreement between her and Beth. So the objection fails.

Partial deference to oneself

We have seen how the equal weight view applies to cases of disagreement with outside advisors. It says that one should defer to an advisor in proportion to one's prior conditional probability that the advisor would be correct. But the view also applies to cases of disagreement with one's future or past self. It constrains how one should take into account the opinions one had, or expects to have, when one has less than perfect trust in one's past or future judgment.

To see how the constraint operates, take the example of being guided by past opinion (the case of being guided by future opinion is similar). Suppose that you have just reevaluated a philosophical theory that you accepted years ago. If not for your past opinion, you would now reject the theory as false. How much should your past acceptance of the theory temper your present rejection? The answer is that you should defer to your past opinion in proportion to your prior conditional probability that your past opinion would be right. More precisely, the crucial question is: setting aside your recent rethinking of the issue, and setting aside your old conclusion about it, which self would you expect to be right in case of this sort of disagreement?

There are a number of factors that determine the answer: whether you think you've gotten sharper over the years, or have received relevant background information, or have caught what is clearly a mistake in your old reasoning. But one factor is of particular note: the degree to which your past and present selves agree on issues surrounding the disputed one.

If there is a great deal of agreement, then you will likely count your past self as a peer, or as nearly one. If so, you should give your past conclusions plenty of weight. In contrast, suppose that between then and now you have undergone a fundamental change in view. Perhaps you have experienced a spiritual revelation, or a quick deconversion. Perhaps you were exposed to a philosophical idea that dramatically reorganized your thinking about a large cluster of issues. Or perhaps you were part of an episode of scientific change that was revolutionary in the sense of Kuhn (1970). In such cases, you will regard your past self as having many mistaken views about issues surrounding certain disputed ones. So in such cases, you should severely discount your previous dissenting opinion.

What this shows is that the equal weight view provides a graded version of Reflection—a version covering cases in which one only partially defers to the judgment of one's future or past self. The view has the effect of tying one's beliefs together over time. But note that the strength of the ties is highly variable. Across periods of ordinary (modest) belief change, the ties are strong. For across such periods, there is a great deal of background agreement. In contrast, across periods in which there is abrupt or dramatic belief change, the ties are weak. For across those periods there is little common ground between the selves at the earlier and later times. In such cases it may be inevitable that one's earlier and later selves regard each other as not to be trusted.

References

Frank Arntzenius. Some problems for conditionalization and reflection. *Journal of Philosophy*, 100(7):356–371, 2003.

David Christensen. Clever bookies and coherent beliefs. *Philosophical Review*, pages 229–247, 1991.

David Christensen. Epistemology of disagreement: the good news. Manuscript, 2004.

Stewart Cohen. Basic knowledge and the problem of easy knowledge. *Philosophy and Phenomenological Research*, 65(2):309–328, 2002.

Richard Feldman. Reasonable religious disagreements. URL `http://www.ling.rochester.edu/~feldman/papers/reasonable%20religious%20disagreements.pdf`. Manuscript, July 2004.

Richard Foley. *Intellectual trust in oneself and others*. Cambridge University Press, Cambridge, 2001.

Haim Gaifman. A theory of higher-order probabilities. In Skyrms and Harper (1988).

Daniel Garber. Old evidence and logical omniscience in Bayesian confirmation theory. In John Earman, editor, *Minnesota studies in the philosophy of science*, volume 10, pages 99–131. University of Minnesota Press, 1983.

Alvin Goldman. *Epistemology and cognition*. Harvard University Press, Cambridge, MA, 1986.

Michael Goldstein. The prevision of a prevision. *Journal of the American Statistical Association*, 78(384):817–819, 1983.

Gary Gutting. *Religious Belief and Religious Skepticism*. University of Notre Dame Press, Notre Dame, 1982.

Alan Hájek. On what conditional probability could not be. *Synthese*, 137:273–323, 2003.

Ned Hall. Correcting the guide to objective chance. *Mind*, 103:505–518, 1994.

Ned Hall. How to set a surprise exam. *Mind*, 108(432):647–703, 1999.

Terry Horgan. Sleeping beauty awakened: New odds at the dawn of the new day. *Analysis*, 64(1):10–21, 2004.

R. Jeffrey. Conditioning, kinematics, and exchangeability. In Skyrms and Harper (1988).

Thomas Kelly. The epistemic significance of disagreement. In John Hawthorne and Tamar Gendler, editors, *Oxford studies in epistemology*, volume 1. Oxford University Press, Oxford, 2005.

Thomas S. Kuhn. *The structure of scientific revolutions*. University of Chicago Press, Chicago, 2nd edition, 1970.

Cliff Landesman, 2000. Personal communication.

Christian List and Philip Pettit. An epistemic free-riding problem? Technical report, RSSS, SPT, ANU, 2003. URL `http://eprints.anu.edu.au/archive/00002233/`.

Dan Moller. Disagreement manifesto. Discussion group notes, 2004.

Philip Pettit. When to defer to the majority—and when not. *Analysis*, 2005. Forthcoming.

Alvin Plantinga. Pluralism: a defense of religious exclusivism. In Philip L. Quinn and Kevin Meeker, editors, *The philosophical challenge of religious diversity*, pages 172–192. Oxford University Press, Oxford, 2000.

Karl R. Popper. *The Logic of Scientific Discovery*. Basic Books, New York, second edition, 1952.

A. Renyi. On a new axiomatic theory of probability. *Acta Math. Acad. Scient. Hungaricae*, 6:285–335, 1955.

M.J. Schervish, T. Seidenfeld, and J.B. Kadane. Stopping to reflect. *Journal of Philosophy*, 101(6):315–322, 2004.

B. Skyrms and William Harper, editors. *Causation, chance, and credence*. Kluwer, Dordecht, 1988.

Brian Skyrms. The value of knowledge. In C. Wade Savage, editor, *Justification, discovery, and the evolution of scientific theories*. University of Minnesota Press, Minneapolis, 1987.

Ernest Sosa. Reflective knowledge in the best circles. *Journal of Philosophy*, 94, 1997.

Robert Stalnaker. The problem of logical omniscience. *Synthese*, 89(3):425–440, 1991.

W. J. Talbott. Two principles of Bayesian epistemology. *Philosophical Studies*, pages 135–150, 1991.

Bas C. van Fraassen. Belief and the will. *Journal of Philosophy*, 81:235–256, 1984.

Bas C. van Fraassen. Belief and the problem of Ulysses and the Sirens. *Philosophical Studies*, 77:7–37, 1995.

Peter van Inwagen. It is wrong, everywhere, always, and for anyone, to believe anything upon insufficient evidence. In Jeff Jordan and Daniel Howard-Snyder, editors, *Faith, Freedom, and Rationality: Philosophy of Religion Today*, pages 137–53. Rowman and Littlefield, London, 1996.

Jonathan Vogel. Reliabilism leveled. *Journal of philosophy*, pages 602–623, 2000.

Jonathan Weisberg. Conditionalization without reflection. Manuscript, 2005.