

Disagreement and Rationality

With saddles for pillows, staring up and contemplating the stars above, two cowboys converse on a warm summer night. Joe says, "What do you think happens when you die?" Billy reflects in silence for a moment and then says, "I think when you die, it's over. There's nothing." Joe ruminates on a stem of grass, top-hand taciturn in demeanor. Finally, "I guess I don't. I don't think it's over. But I don't know."

The conversation is intriguing philosophically if we take the assertions as accurately reflecting what internal states are present. Joe apparently believes things that he believes he doesn't know, but more relevant for our present topic is that Joe and Billy disagree. Joe thinks that in some sense we survive death; Billy doesn't.

Some philosophers will feel the need to defend the rationality of Joe and Billy in some way. They will hold that the only way for both Joe and Billy to rationally believe what they do is for further factors to be present. Maybe Joe has some evidence that Billy doesn't, or vice-versa; maybe one is a dolt and the other the next Einstein. For such philosophers, disagreement always delivers the scent of irrationality, a scent cleansed from the epistemic air only by the presence of factors such as those just noted. But, for such philosophers, once we control for a comparatively short list of ways to explain away the disagreement, epistemic irrationality is unavoidable.

I doubt, however, that after reading the introductory paragraph, any scent of irrationality was detected. The scene is completely innocuous from an epistemic point of view. Describing the two as cowboys conveys some sense that there is no vast difference between the two in terms of intellectual competence, nor are cowboys noted for having a need to think of themselves as

intellectually superior to their companions, and the subject matter is selected to block any tendency to think that one of them has special information about the subject that the other doesn't. They disagree without rancor, and there is no tendency on our part to sniff to find some purported irrationality in the story. There is nothing for the irrationality sleuthhounds to track here.

There is a caveat, though, since some philosophers will think that beliefs about such matters are all automatically irrational, and so might explain away the absence of an epistemic response to the story. If so, we can change the example, since the general point remains that even when we change the subject matter about which disputants disagree, sleuthhounds still have no scent to follow, whether the subject matter is the relatively sophisticated matters of global warming, voucher programs for private schools, capital punishment, or the more mundane matters of whether it will rain tomorrow, which schools and doctors and dentists are best for our children, or which sports franchises have the most illustrious pasts. Disputation and disagreement abound at every turn; suspicions of irrationality are rare.

My goal here is to provide the epistemological underpinning to support this view of disagreement, the view according to which rational disagreement is unproblematic. I'll begin with some problems for the alternative view. These problems show the need for a restricted view of the significance of disagreement, and I will develop three models for such restrictions. We will see that though disagreement has some epistemic effect, it doesn't have the sort of effects the irrationality hounds think they detect.

I. Mollificationism and Its Discontents

Let us call the position that views disagreement as a sure sign of irrationality “mollificationism”. Mollificationists see harmony as the equilibrium point in the theory of rationality. People should agree in attitude, and when they don’t, there is a disturbance in the force that needs attention.

Mollificationism faces serious problems, however. Notice, first, what happens when one insists that agreement in attitude is the only natural point of equilibrium in the theory of rationality. Mollificationists like to attend to cases of disagreement in which one person believes p and the other $\sim p$, but this is only one kind of attitudinal distance between people. There are, at the coarse-grained level, four possibilities: believing, believing the opposite (disbelieving), withholding, and taking no attitude at all. And if we move to the fine-grained level of degree of belief, there are as many possibilities as there are real numbers between 0 and 1.

Mollificationists hold that differences in attitude are always and everywhere rationally suspicious, and once we see the variety of attitudes possible, this view should look strange and bizarre. Even if we restrict ourselves to the coarse-grained level, the mollificationist has strange demands. If you and I disagree about p , it is not enough that each of us gives up our beliefs. If you give up your belief and come to withhold concerning p , but I give up any attitude at all because I don’t know what attitude fits my evidence any longer, the mollificationist is still unhappy. We still disagree in attitude: you withhold and I take no attitude at all toward p . What are we poor creatures, desirous of being rational, to do now? One of us must defer to the other, but we better not both do it at once! The same issue arises if you withhold and I believe. One

must defer to the other, but which to which? Such questions push mollificationism in the direction of default agnosticism, the position that says agreement in attitude is required and the natural equilibrium point for such agreement is that of withholding about any matter in dispute. It is hard to see how such a position is compelling, as if the best advice we could have received when starting our intellectual journey would have been, “Always remember to prefer knowing nothing at all to being disagreeable.” One can hear in the distance the skeptic rejoicing. The route isn’t as direct as one gets from the Jamesian contrast that leads to withholding for fear of being duped,¹ but instead along a different social path instead, a path that would fit quite well with a preoccupation with being able to get along with everybody. At bottom, however, both paths lead to the same skeptical end, and once we nonskeptics see this point, we will endorse mollificationism only if it can be articulated in a way that doesn’t land us in skepticism simply in virtue of the fact that there are skeptics around.

The fundamental point here, however, is not one about avoiding skepticism, but about the paralyzing effect that comes from thinking of having the same attitude (or lack of one) as the equilibrium point for rationality. In some cases, no movement to an intermediate stance is possible for both individuals: all that is possible is for one of the two to adopt the stance of the other. When one person takes no attitude and the other person withholds, there is no intermediate position available.

The mollificationist might respond that there is no special problem here, that this is simply how compromises work in some extreme cases. In lots of cases, there is an intermediate position available, but in other cases there isn’t; in both kinds of cases, however, the demands of rationality still impose a requirement of compromise in the face of disagreement in attitude.

The language of compromise is rhetorically useful here for the mollificationist, and it is important to see why it is inappropriate language to use to describe the view in question. To see why, consider first how ordinary compromises occur and what they involve. Suppose two politicians disagree about policy. In the end, some resolution is necessary, since unending paralysis is intolerable (for whatever reason). So they compromise. Both think the result is less than ideal. To understand the result, we need to know not only the history of the process, what they used to think and why, but also what they presently think and why. They think that the compromise is best, in some sense, but also, in another sense, that it is not. Without some such internal conflict, we don't yet understand the political process in question or the full nature of the accommodations that have been made by one or both parties. Compromise requires some such internal conflict, whereas had one of the two convinced the other of the correctness of his or her view, no such conflict would be present and no compromise would have occurred.

Mollificationism can be pleasingly put in terms of the language of compromise, involving the claim that when two cognizers disagree about some claim p , there is rational pressure for compromise. Such language is misleading, however, since situations of cognitive compromise are different from situations of full resolution of disagreement. When resolution occurs, the story we tell has present agreement as the outcome of past disagreement. When cognitive compromise occurs, we expect something akin to what we found in the political case. As the disputants become more aware that they are converging on a point where rationality compels them to abandon their beliefs, they will view the approaching event with consternation. They will view it as an intellectual loss to mollify their attitude, and this sense of loss will not leave once the convergence point is reached. Just as in the case of political compromise, some present mark

will remain in place to distinguish it from cases of epistemic resolution. It is not enough to note only that the two used to disagree and now no longer do, since that doesn't distinguish compromise from mere change in view.

It is difficult to see what decent answer the mollificationist can provide here to explain the difference. After compromise, the two now agree about p : they both withhold judgment, let us suppose. How did this come about, in a way that makes compromise different from mere change in view? The temptation is to say that they assess the force of the evidence differently. But this can't be the story if rationality requires compromise, since then X and Y still disagree: one's assessment is that the evidence supports p and the other that it supports *not- p* . In order to be rational, the two will have to give up this view as well. If all we can say is that they used to disagree but no longer do, we have no difference at all between compromise and mere change in view.

Perhaps, though, the mollificationist can go internal here. Each can hold that from their own point of view the claim in question appears true or false, respectively, and that appearance doesn't disappear when compromise occurs. Putting the point in this way threatens to undermine Mollificationism almost immediately, however, for rationality is perspectival in just the sort of way that, if the two points of view differ in the way imagined, then a difference in rational attitude is to be expected. What the Mollificationist needs here is a distinction between total point of view (on which rationality supervenes) and some more partial point of view on which differing appearances remain in place. Such a distinction, however, is present in cases of fully resolved disagreement as well, however, since when you convince me that I've made a mistake, I can still see the aspects of the situation that led to my mistake. So the difficulty of distinguishing

a situation of compromise from one involving a more ordinary resolution of disagreement through change in view remains, and I see no way to defend the use of the language of compromise here.

There is a more direct complaint as well to lodge against the idea that compromise is the default order of the day in the face of disagreement. The more direct complaint concerns what I will call cognitive self-alienation. To see what cognitive self-alienation involves, let me begin indirectly. Each of us has certain cognitive abilities and disabilities. Some of us are better at math, some are better at noticing small details, some are better at visual detection, etc. In addition to these differences in cognitive abilities, there is also the matter of our own view of ourselves as to the level of ability we have in a given domain. This perspective on self can lead us to demur on changing opinion in the face of disagreement and can also lead us to defer to others in certain circumstances, such as when we view them as in a better position on the matter. Whether to demur or defer is a matter of our perspective on ourselves.

In between demurring and deferring is desisting in belief, which is perfectly sensible from each of our points of view when our view of ourselves falls between a view calling for one to demur or for one to defer. To defer and to desist both involve change in view, but all three responses are here understood in terms of the relationship between a first-order response to disagreement and a metalevel perspective on oneself that makes sense of the various responses of demurring, deferring, or desisting.

The notion of compromise championed by mollificationists, however, is different. Recall that we are supposing that compromise is required in such a way that points of view can differ in the way imagined between X and Y (so that from X's point of view p is true and from Y's $\sim p$ is

true). Since resolution of disagreement has failed by ordinary means, each is aware of the presence of someone who disagrees, and neither has a perspective on self that removes the disagreement. We thus have individuals involved in a disagreement whose views of self are incompatible with desisting or deferring, but rather calls for demurring on the issue of changing opinion. And yet, compromise is required, according to the Mollificationist. Such a requirement, however, insists that the perspective on self taken by the two individuals be abandoned. Neither is allowed to take the demurring perspective on self, no matter how they came to such a view of self, and instead must take the attitude of desisting or deferring. The mollificationist position thus requires cognitive self-alienation, where one has a view of oneself and one's abilities that coheres fully with all of one's prior experiences and one's total conception of things, and yet which cannot rationally play a role in what attitude is legitimate at the first-order level. Instead of a happy union and coherence between first-order belief and meta-level attitude toward self, we have alienation and cognitive hostility between these levels required by the theory of rationality in question. A theory that insists that rationality requires such cognitive self-alienation has a serious burden of explaining how the theory is appropriately sensitive to the perspectival character of rationality.

So the problem of cognitive self-alienation is a problem of failing to honor appropriately the obvious perspectival character of rationality. The mollificationist emphasis on compromise requires discounting aspects of a perspective in determining what attitudes are rational. In any case of disagreement, part of what needs to be assessed is whether attitudes of deferring, desisting, or demurring are most appropriate. The mere fact that someone has one of the three attitudes doesn't by itself make that attitude appropriate, but which attitude is appropriate should

be a matter determined from the point of the view of the perspective of the individual in question. That is the lesson of the perspectival character of rationality, and the implication of it is that the mollificationist demand for compromise will often require cognitive self-alienation because the perspective of the person in question makes demurring the appropriate attitude when the mollificationist insists on desisting or deferring.

The problem of cognitive self-alienation is not the only problem that mollificationism faces, however. Another problem for the view is its tendency toward self-defeat.² To see the problem, begin with the hyperbolic thesis that whenever people disagree, they are both irrational. Since I think that view is obviously false, no one can rationally believe the hyperbolic thesis. It is a philosophical thesis undermined by a quite prosaic fact, the fact that some of us think that disagreement is at least as expected in matters philosophical as in any other. Moreover, consider the implications of this epistemological hyperbole. When tempted to such a view, the first thing a responsible cognizer should do is check to see if others disagree. When we can and the search is easy, we should check for such a pearl of great price. And it will be easy to find. E-mail a few friends about anything in philosophy and you will have it. But then not only the hyperbolic thesis itself cannot be rationally believed, neither can any disputed thesis in any part of philosophy or elsewhere.

We should not label the horribly hyperbolic thesis “self-defeating,” however, at least not necessarily so. It isn’t, since it is possible that everyone agrees with the thesis. A nice analogy here is with two versions of the Liar Paradox. The direct version of the paradox occurs when we have a sentence such as “this sentence is false.” But there are indirect versions, such as when a shirt has the following sentences on front and back, respectively: “the sentence on the other side

of this shirt is false,” and “the sentence on the other side of this shirt is true.” Here we have a kind of contingent undermining of each sentence. If we want to say that each sentence is a liar sentence, defeating its own truth, we will have to say the self-defeat is contingent rather than necessary as in the former case. Perhaps the language of self-defeat isn’t the right language here, but terminology isn’t really the point. What matters in both cases, semantic and epistemic, is that refutation occurs on the basis of contingent factors that are in place. In the semantic case, it is the sentence itself that is defeated; in the epistemic case, it is belief in the hyperbolic thesis the rationality of which is defeated. In the epistemic case, what matters is that these contingent factors are so obvious and predictable that one needs a really serious epistemic blindspot not to take such a factor into account before endorsing the horribly hyperbolic position. We may want to say that the position is contingently self-defeating in the way the shirt sentences are, and I will adopt that terminology here, with the forewarning that contingent self-defeat may not be a form of self-defeat at all together with the additional proviso that the object of defeat is the truth of a sentence in one case and the rationality of belief in the other. What matters most is the clear point that the flaw in question is debilitating.

Contingent self-defeat doesn’t only apply to the hyperbolic position, however. For those who detect a scent of irrationality and try to confirm it by additional factors, the same problem remains. If you think that scent and reality converge when the disputants are epistemic peers, the same problem remains; if you think such convergence occurs when epistemic peers share all and only the same direct evidence, the embarrassing problem is still present. I will not offer my own e-mail address for proof, but will now relativize a bit, in case I’m not your peer or your knowledge of epistemology exceeds mine. The thesis may not be contingently self-defeating for

absolutely everyone, since there may be an epistemologist who is either smarter than all the rest or who has better information than all the rest. So the claim could be believed by the smartest and best informed, but by no one else. For everyone else, though, the thesis is so likely to be contingently self-defeating that is no better off than the horrible hyperbole itself.

Perhaps we could put the point this way. There are too many assumptions needed for the amended hyperbole to be adopted. One assumption is that you are smarter and better informed than anyone who disagrees with you about the amended hyperbole. But there is another assumption that even the most arrogant among us will have trouble with. There will be lots of epistemologists in the future. Many of them will be really smart. Many of them will know lots more about epistemology than we do. And, I would bet, some of these incredibly smart and well-informed epistemologists will think the amended hyperbole is false. But maybe I'm wrong. The point, however, is that it is an assumption needed for the amended hyperbole to avoid contingent self-defeat that I'm wrong. No one should be so intellectually arrogant as to commit themselves to the view that there will never be a brighter and better-informed disputant of the amended hyperbole.

Moreover, going down the hyperbolic path should seem, once we begin investigating it, a completely wrongheaded way to proceed intellectually. We want to get to the truth and avoid error. How to do so is often not clear, but it is easy to describe things at a very abstract level in terms of the concept of evidence or indications of truth or signs of truth. We seek indications of truth and falsity, and adjust our views to such signs. We want to be the kind of individual who looks for evidence of truth and follows the evidence where it leads. Nowhere in this story of the intellectual life does our neighbor enter in. When we want to know whether p is true, we don't

turn to surveys to find if there are people who think p is false. Because we don't, we don't in turn try to sort these naysayers into more or less intelligent, more or less informed. At best, other people are repositories of the evidence we seek, not a source of basic evidence itself. That is how we behave when seeking the truth, and any account of the epistemic significance of disagreement had better take account of this practice.

One might object that this description is woefully inadequate because it leaves out the role of testimony. In the course of finding the truth and avoiding error, our neighbor plays a key role as a source of information. This point is correct, but there is a quite intuitive story to tell about the difference between testimony on the one hand and the issue of disagreement on the other. Testimony is important because others are repositories of evidence (or else there is a chain of testifiers leading back to such a repository). Once we learn that the person speaking has no information about the matter beyond what we also have, that person's word ceases to have the ordinary power of testimony, which the word of others has in virtue of our lack of information on the subject, and becomes an expression of disagreement. There is no need to treat disagreement in precisely the same way we treat testimony in order to provide a proper place for testimony in our account of how we develop a rational understanding of a given subject matter.

There are complications that would need to be addressed in a full discussion of the relationship between the epistemology of testimony and the epistemology of agreement and disagreement, but my point here is simple enough that we can bypass the complications at present. We are considering the self-defeating character of simple views on which disagreement undermines rationality. In response, I pointed out that we don't engage in inquiry by taking surveys as seriously as such views would require, and the response to this simple point was that

this simple point ignores the importance of testimony. My response is that we don't need to place the kind of importance on opinion surveys that the self-defeating views imply in order to take testimony seriously in our epistemology. Refining this point to make it precise would be an interesting detour here, but the point is so obvious that no such detour is needed to appreciate it.

II. Modeling the Epistemic Significance of Disagreement

All of this seems so obvious and compelling that one wonders how one could favor the view that disagreement imperils rationality at all. There is a legitimate worry here, however, and it arises from the fact that sometimes we do reasonably abandon a view because we find others who disagree. We need an explanation of this phenomenon in order to make sense of the significance of disagreement.

We can begin to see through the thicket of issues by examining a relatively simple case in which we come to trust the readings of an instrument when it conflicts with our own judgement. The ideal case of such deference is a calculator that calculates from axioms that we know to be true. Humility regarding our own capacity for errors in arithmetic lead us to defer to the machine when we have a conflict, and the reason is that the expression of humility appropriate to us is not appropriate regarding the machine. We also defer to machines when the amount of information they detect exceeds ours. We don't have eyes in the back of our heads, but machines can be built that do. As a result, they may know what is behind us and we will defer to them in such a situation because of their superior capacities. So there are two ways in which deference to another can arise legitimately: when they have more information than we do and when they are

better analysts of the information than we are.³

We can model the epistemic effects of disagreement in three different ways. One way is to take opinions of others as evidence regarding the proposition in question. Another way is to treat the opinion of another as a way in which people typically signal that they have information that shows that the claim in question is true. On such a model, the word of another isn't immediately and directly evidence for the claim in question, but rather evidence that there is evidence for the claim in question.⁴ If we add to this model some account of how to detach the evidence operator (so that one goes from "there is evidence that there is evidence for p" to "there is evidence for p"), then the model tells us conditions under which the opinion of another is evidence for the claim in question, though only indirectly so.

The difference between these two models concerns whether the word of another is evidence or meta-evidence, but once we notice the option of moving to the meta-level, there is a third model that might be endorsed. In the theory of defeat in epistemology, the usual view is that of John Pollock according to which there are two fundamental types of defeaters: rebutters and undercutters. A different model might distinguish between testimony and disagreement, claiming that testimony must be modeled by one of the first two models just noted (either evidence or meta-evidence for p), but that disagreement must be modeled in terms of defeat. If disagreement is modeled in this way as a rebutting defeater, it is a natural accompaniment to the view that testimony provides evidence for the claim in question, for a rebutting defeater of my evidence for p is evidence against p. Thus disagreement is just a special case of testimony.

The alternative view is that disagreement can function as an undercutting defeater. An undercutter for my evidence for p is a claim that supports the view that the evidence isn't a

reliable indicator of the truth of *p*. For example, suppose you know that if you seem to see a pink elephant, your eyes can't be trusted. This information functions as an undercutting defeater for the usual evidential support relation between seeming to see a pink elephant and the claim that there is a pink elephant nearby.

For present purposes, it is not necessary to single out one of these models in preference to the others, but there are a couple points worth noting about them. First, the defeater model faces a rather serious objection, since it is forced to treat disagreement as different from testimony more generally conceived. There may be differences between the epistemic significance of disagreement and testimony, but it would be surprising to find them functioning in such a different way so that one model was appropriate for testimony (either the evidence model or the metaevidence model) and a different model appropriate for disagreement (the undercutting defeater model).

Second, two quite different factors pull us in opposite directions when considering the remaining models, on the assumption that we want to model disagreement and testimony in the same way. Consider first what is involved in using a measuring device as a kind of testimonial instrument. Imagine that you want to know the range of light conditions shining on a particular window in your home, from sunup to sunset. You don't have time to watch all day, and besides, if you look down to record an observation, you'll miss some information. So you want to build a detection device that records the observations automatically.

Of course, you don't just construct the thing and then take it's word for the truth (it generates, let's say, a continuous graph plotting light wave lengths against time). You have to calibrate it to make sure that it is accurate. So you do that. Prior to calibrating it, it didn't

provide undefeated evidence, if any evidence at all, about the matter in question. Prior to calibrating it, you had no information one way or the other about whether the machine was reliable. You need reason to trust it to use it to acquire information about the target of inquiry. Now, failing to have any such meta-information is not itself a defeater of the confirming power of first-order information: I don't have to have information that I'm reliable before acquiring reasonable beliefs on the basis of inquiry. Moreover, calibrating the machine doesn't change its capacities with respect to my inquiry: it doesn't somehow turn the machine into a first-order evidence generator when it was only a second-order evidence generator before. Clearly, prior to calibration, the graph generated is evidence that this machine is responding differentially to different lighting conditions. We have evidence that the device is responding to information about the target of inquiry; it is a relay device, relaying information about something I want to know about. The information relayed is, of course, evidence: it is the information relevant to the question of what to believe about the target of inquiry. So the machine's activity is best modeled as a metaevidence device: its readings are evidence about the evidence regarding the target of inquiry.

Viewed in this way, we have a nice story to tell about the machine before and after calibration. Prior to calibrating the device, we aren't in a position to detach the evidential operator: we don't have a system of information that allows us to conclude that there is evidence for a given claim on the basis of the machine readings which give us evidence that there is such evidence. Prior to calibrating the device, we can't rely on the machine in arriving at reasonable beliefs about the target of inquiry, but after calibration, such a possibility exists, and the difference between these two situations is just the difference between only having evidence that

we have evidence for the claim in question and actually having evidence for that claim.

Such a picture may seem to conflict with the basic trust necessary for early learning and the rationality involved in it. Much of what we know depends on taking the word of others prior to having anything like the information generated by the calibration process described above for measuring devices. In being rational animals, we come equipped with default cognitive mechanisms and as learning progresses we come to adapt these default mechanisms in various ways, including the development of wariness about sources of information. So the description of calibrating a detection advice is relevant once a certain level of sophistication is achieved, but it can't be the basic description that applies to all testimony.

Once we see this point, we might wonder about its implications for the two remaining models under consideration. On the metaevidence model, there are two stages needed to use a source of information as a guide to belief. In the first stage, one must detach the metaevidence operator to get the information that there is evidence for the claim in question, and in the second stage, one must detach this remaining evidence operator to conclude that the claim in question is true. On the evidence model, only one detachment step is needed.

Once seen in this way, there is some pressure to view the metaevidence model as an overly intellectual model of how learning by testimony occurs. Take your favorite three-year-old, and consider what it means to say that this small child is detaching an evidence operator twice over. On the evidence model, the child just hears what is said and detachment is just a matter of taking the word of the person in question: it is to come to believe the claim that is asserted by the source. But on the metaevidence model, there must be a first detachment involving the conclusion that there is evidence for the claim in question, and surely it is

implausible to view this child as forming beliefs about evidence prior to forming beliefs about what is asserted. So the metaevidence model seems to over-intellectual the process of belief formation on the basis of testimony.

I think this criticism misconstrues the metaevidence model, however. On this model, in order for testimony to make a belief rational, an intermediate step is required in which one moves from having evidence of evidence for p to having evidence for p . But the intermediate conclusion need not be present in the cognitive system in question in the form of belief. The most common approach to detachment rules begins from the idea that high probability, or a high degree of evidential support, is not sufficient for rational belief, but it is in the absence of defeaters.⁵ Defenders of the metaevidence model might transpose this idea into the present context by claiming that the presence of evidence that there is evidence for p is itself evidence for p when there are no defeaters present to undermine the connection between the claim about metaevidence and the claim about evidence. Thus, when the small child hears a parent say, “There are cookies on the table,” evidence exists in the form of an experiential state in the child. This state, on the metalevel model, is initially only metaevidence, but in the absence of defeat, it is also evidence; and once we reach the point that it is evidence, it makes the belief of the child rational so long as there are no defeaters to undermine the connection between it and p itself. Given this approach, two types of defeaters must be absent, one kind concerning the relationship between the existence of metaevidence and the existence of evidence itself, and the other kind concerning the relationship between the existence of evidence and the object of belief.

The key to avoiding the overintellectualization charge, then, is to refuse to treat the model in terms of inference structures that require an additional piece of information in the head of the

child before allowing rational belief in the claim in question. Instead, the evidence claims themselves are cast in terms of properties of the state itself that is evidence. Thought of in this way, the experiential state in question has the property of being evidence for p always and everywhere on the evidence model, but it has this property on the metaevidence model only when there are no other pieces of information internal to the agent in question that defeat the supports relation between the state in question and the claim that this state is evidence for the target proposition in question. This supports relation, however, is not a relation between an experience and some further belief or experience, but is simply a relation between the experience and a proposition.

If such a story can be sustained, we can model sophisticated reliance on instruments that require calibration in the same way we model early learning by testimony, and that result would be pleasing indeed. Moreover, there is a further point to make on behalf of this way of conceiving of testimony deriving from the lessons about levels confusions taught to us by William Alston. Alston argued that it is not a good idea to let the skeptic use second-order claims to undermine the epistemic status of first-order claims.⁶ For example, when the skeptic asks how I know that it is raining, and I cite some evidence for it, the skeptic may question how I know that this evidence is a reliable indicator of rain. That's an interesting question, but one that is irrelevant to whether I know that it is raining. It may be relevant to the question of whether I know that I know that it is raining, but if we hold that evidence needs to reliably indicate in order to generate knowledge, we do not also need to hold that we must know that evidence is a reliable indicator in order to know. Keeping clear of such levels confusions helps avoid a very quick argument for skepticism, since the skeptic can always ask about the metalevel $n+1$ when we

answer his or her query about level n . In the same way, a defender of the metaevidence model may insist that the state of information that is the basis for a rational belief must also provide a suitable basis for the claim that it is evidence for the claim in question without requiring that one believe or rationally believe this further claim. Its truth is required, but not any belief or awareness of this truth.

II. Testimony and Disagreement

As noted earlier, however, we should expect not only similarities between a theory of testimony and a theory of disagreement, but also some account of the differences between the two. The models just considered give us the resources for explaining the similarities, and they also help us understand how the two theories will diverge as well. We can see how by considering a mollificationist claim about a particular disagreement. Consider David Christensen's discussion of such a case in which he and his friend are dividing up their respective shares of a bill at a restaurant, and they come up with different answers:

Given that my friend and I are generally reliable thinkers who have studied the same evidence, the fact that we disagree will be explained by the fact that at least one of us has made a mistake in the case. But intuitively, the explanation in terms of my friend's mistake is no more reasonable than the explanation in terms of my mistake. And I should acknowledge this by moving my belief toward hers.⁷

The argument here begins by assuming that both parties are equally competent and have the same evidence (and, we may assume, both reasonably view the matter in this way as well), and

concludes that both parties are required to abandon belief in the conclusions they reached. The key claim in the example involves the idea of explaining the disagreement. There are two ways in which the disagreement can be explained: either in terms of one person having made a mistake or in terms of the other person having made a mistake. The key claim here is that neither explanation is more reasonable to adopt than the other.

The idea here, I take it, is that if your information includes parity of evidence and parity of competence regarding one with whom one knows that one disagrees, then change in view is required. The passage itself actually employs the stronger position that the mere fact of parity itself requires change in view. This stronger position is obviously flawed, however, since the facts may be as stated and yet it be rational to deny them (to say nothing of being unaware of the disagreement itself). So we should focus on the weaker position, that includes among your background evidence the parity in question and the disagreement itself.

The example is well-chosen to make plausible the principle enunciated, that parity information plus knowledge of disagreement requires attitude adjustment. The example, however, does support mollificationism, interpreted in terms of requiring compromise of attitude even when one's view of self is in opposition to demurring. In this case, demurring is the natural attitude assumed to be present in each party, and hence the case is not a case of intellectual compromise at all. But because the example makes plausible the principle that parity information plus knowledge of disagreement requires attitude adjustment, it provides a good example of how to separate the theory of testimony from the theory of disagreement. The difference concerns the defeasible character of epistemic support (so that the notion of requirement in the principle just stated is interpreted in a defeasible fashion). Testimony

provides reasons for belief because, in a wide variety of cases, no internal defeaters are present concerning the evidentiary power of testimony. But in cases of disagreement, the default position is that the power of testimony is blocked because of the present of a defeater: namely, the belief itself and whatever evidence one has for it! Now defeaters are themselves capable of being overridden by further information, so the mere fact of belief doesn't imply that disagreement never forces revision of belief. Moreover, awareness of disagreement, on either the evidence model or the metaevidence model, will often itself be a rebutting defeater for the supports relation between one's evidence and the belief in question. So we have a case with defeaters all over the place, and some reflective sorting is in order to determine which defeat relationships take priority.

It is here that one's perspective on oneself as a cognitive being provides an arbiter in the form of an overrider in one direction or the other. That is the lesson of the discussion of deferring, demurring, and desisting. Which of these options is the rational one to adopt depends on the degree of self-trust that remains about the specific issue in question. As self-trust dissipates, desisting and deferring become more rational in virtue of the loss of an overrider for the defeating role that disagreement plays in rational cognitive life. But when self-trust remains, there is an overrider for the defeating role that disagreement by those recognized to be peers or even superiors in terms of data and analysis of data.

One might question what to say when the self-trust in question is itself clearly irrational. In such a case, there will be other elements in the system of information that imply this irrationality, and then, of course, the overriding power of the level of self-trust is itself overridden

in turn. In the absence of such further information, however, the level of self-trust is the final arbiter in cases of disagreement.

The upshot of this account of the difference between testimony and disagreement is insight into the degree to which rationality is a subjective matter. As a person loses trust in self as a cognitive being, rationality dissipates, but this is just as it should be. When self-trust in an area is lost, the views that used to be reasonable to hold in that area are no longer reasonable, and so retaining them unchanged would be irrational.

The role for self-trust here is limited. There is no claim here to the effect that in every instance of rational belief, there is an attitude of self-trust that plays a role in explaining the rationality in question. The attitude of self-trust arises in cases of disagreement because of the nature of the case in question, not because of the nature of rationality. When we encounter someone who disagrees, that information constitutes a defeater that affects the rationality of our belief. Exactly how it affects this rationality depends on which model of testimony and disagreement is best, but let me answer here from the perspective of the metaevidence model. On that model, we have evidence of a rebutting defeater, and thus will have a rebutting defeater unless there is a defeater present to block the support in question. It is here that questions of data and competence are relevant, since disagreement by someone with inferior information or inferior abilities to analyze information would be just such a defeater. So learning about peer status in these respects serves to undermine potential or actual defeaters here. My inclination here is to claim that known ignorance of peer status is a defeater of the evidence for a rebutting defeater. In any case, the cognizer in question has to turn reflective in order to avoid the negative rational effects disagreement portends. Once having turned reflective about peer status, the

question of self-trust in the domain in question arises automatically and thus becomes relevant.

In non-reflective cases where the issue is not raised, nothing said here suggests that some attitude of self-trust plays a role in the story of rationality, though of course nothing said here is incompatible with that view either.⁸

Endnotes

1. The exact quote from William James's "The Will to Believe" is: "he who says, "Better go without belief forever than believe a lie!" merely shows his own preponderant private horror of becoming a dupe."
2. I believe the first printed version of such an argument against a version of mollificationism is in Plantinga's 1995 piece entitled "Pluralism: A Defense of Religious Exclusivism." Peter van Inwagen's piece "Is It Wrong Everywhere, Always, and for Anyone to Believe Anything on Insufficient Evidence?" contains such an argument as well, and was presented at the Chapel Hill Colloquium in 1993, but it wasn't published until 1996. It is instructive to note that defenses of mollificationism since these pieces were published do not address the argument from self-defeat.
3. See Ned Hall, Adam Elga, Thomas Kelly, and James Joyce for expressions of these distinct types of deference. Ned Hall. (2004) "Two Mistakes About Credence and Chance", *Australasian Journal of Philosophy* **82**: 93-111; Adam Elga. (forthcoming in *Noûs*) "Reflection and Disagreement"; Thomas Kelly. (2005) "The Epistemic Significance of Disagreement", *Oxford Studies in Epistemology* **1**: 167-96; James Joyce, "Epistemic Deference," *Proceedings of the Aristotelian Society* (forthcoming).
4. Thomas Kelly, "The Epistemic Significance of Disagreement." *Oxford Studies in Epistemology*, Edited by John Hawthorne and Tamar Gendler, Oxford: Oxford University Press, 2005, pp. 167-196.
5. For discussion, see Timothy Williamson and Igor Douven, 'Generalizing the Lottery Paradox', *British Journal for the Philosophy of Science*, 57, 4 (2006): 755-779.
6. William P. Alston, "Levels Confusions in Epistemology," *Midwest Studies in Epistemology*

volume 5, 1980, pp. 135-150.

7. David Christensen, "Epistemology of Disagreement: The Good News," *The Philosophical Review* 116 (2007), pp. 187-217.

8. For a defense and explication of a view that requires a stronger role for self-trust, see Keith Lehrer, *Self Trust: A Study of Reason, Knowledge and Autonomy*, (Oxford: Oxford University Press, 1997).