

VEILED DISAGREEMENT*

How I should weigh my disagreement with you depends at least in part on how reliable I take you to be. But comparisons of reliability are tricky: people seem reliable and turn out not to be; they prove reliable in some ways but not others. A theory of how rationally to respond to disagreement requires a clear account of how to measure comparative reliability. Here I show how such an account can be had by drawing on the contractualist strategy of reaching moral and political agreement by imposing restrictions on the information available to disputants. A rational response to disagreement requires considering all the particular details of the dispute at hand, but behind a veil of ignorance that precludes awareness of one's own position in the debate. Imposing the right sort of veil resolves several of the leading puzzles that confront existing theories of disagreement. It also sheds an interesting light on the very different ways in which disagreement gets resolved in epistemology versus political theory, raising troubling questions for both fields.

I. THE GENERALITY PROBLEM

The most widely discussed thesis concerning disagreement—and the thesis that I will defend—insists that rationality requires an attitude of impartiality: giving equal consideration to one's opponent and oneself. Equal consideration for the views of others does not always require equal credence: if I judge myself more reliable than you, then I should give more credence to my views. But impartiality requires not favoring my own views just because they are mine, or just because they seem true to me. Impartiality entails that when two seemingly equally reliable agents disagree in some domain where they are equally well informed, the rational course of action is for each agent to give equal weight to the other's view. It is easy to see how rationality might seem to require this, but also easy to see that the consequences of such a policy would be startling, inasmuch as there seem to be many everyday circumstances in life (religion, politics, philosophy, and so on) where we are disposed to maintain our beliefs even in the face of intelligent, well-informed opposition.

*I owe thanks for their help to Adam Hosein, Alison Jaggar, Bradley Monton, Michael Tooley, the editors of this JOURNAL, and to an audience at the CU-Boulder Center for Values and Social Policy.

Let us refer to this equal-consideration doctrine as the thesis of *Impartiality*. Since the *prima facie* plausibility of Impartiality as a principle of epistemic rationality is obvious, and has been argued for in detail by others, I will take it for granted here as my starting point.¹ The thesis is, however, highly controversial, despite its intuitive plausibility, because it is not clear that the consequences of Impartiality are ones we can live with. Part of what I seek to show, then, is why some of the worst apparent consequences of the thesis do not in fact arise.

The most discussed consequence of Impartiality is that in the special case of *peer disagreement* we should give equal weight to the views of our epistemic peers, and accordingly suspend our beliefs. Such an occurrence depends crucially on various details of how the situation is set out: one's opponents must be equinumerous with oneself and one's allies, the two sides must have equal and opposite confidence regarding the proposition in question, all parties must share the same information, and be equally reliable. The point is simply that Impartiality makes it rational to suspend belief in the face of disagreement only if the circumstances of the case are set out quite carefully and idealistically. It may accordingly be questioned whether Impartiality will actually make much difference in the real world—whether the allegedly startling consequences will ever actually obtain.

Of the various idealizations that Impartiality requires, most problematic is the demand of equal reliability: that the disagreement concern *peers*. Here the complaint has been often expressed that it is not even remotely realistic to suppose that two agents will be equally reliable cognitive agents, let alone that we could ever have any good basis for supposing someone else to be our peer in this way.² This

¹ Prominent statements of the Impartiality thesis can be found in Richard Feldman, "Epistemological Puzzles about Disagreement," in Stephen Hetherington, ed., *Epistemology Futures* (New York: Oxford, 2006), pp. 216–36; and in Adam Elga, "Reflection and Disagreement," *Noûs*, xli, 3 (September 2007): 478–502. Elga there coins the often-used but potentially misleading label 'equal-weight view'. Compare David Christensen, "Disagreement as Evidence: The Epistemology of Controversy," *Philosophy Compass*, iv, 5 (September 2009): 756–67, who prefers to speak of "conciliatory" views. The term 'conformism' is favored by Jennifer Lackey, "A Justificationist View of Disagreement's Epistemic Significance," in Adrian Haddock, Alan Millar, and Duncan Pritchard, eds., *Social Epistemology* (New York: Oxford, 2010), pp. 298–325. But she herself does not wholly endorse the view, for reasons considered below. The term 'Impartiality' is my own.

² For worries of this sort see Richard Feldman, "Evidentialism, Higher-Order Evidence, and Disagreement," *Episteme*, vi, 3 (October 2009): 294–312, at pp. 300–01; Bryan Frances, "The Reflective Epistemic Renegade," *Philosophy and Phenomenological Research*, lxxxi, 2 (September 2010): 419–63; and, in detail, Nathan L. King, "Disagreement: What's the Problem? or A Good Peer is Hard to Find," *Philosophy and Phenomenological Research*, lxxxv, 2 (September 2012): 249–72. Here I limit the concept of *epistemic peer*

line of objection can be quickly blocked, however, by expressing the theory in terms not of known reliability but of expected reliability.³ On this construal, you calculate your own expected reliability as a function of how likely you take it to be that you are reliable to this degree or that degree, and you calculate the reliability of your opponent in the same way. There is no need to suppose yourself and others to be exactly equally reliable; there is no need to *know* (or even have good reason for believing) just how reliable you and others are. It is enough to generate the problem of peer disagreement for you to be in a situation where you have no good reason to think yourself any more or less reliable than your opponent. Such situations seem common enough.

My concern will be with a deeper worry about equal reliability, a worry over scope. An agent's cognitive reliability will vary over differences in subject matter, information, and conditions. We are all better or worse at getting at the truth in some domains than others, and so estimates of our reliability ought to vary widely, depending on how broadly or narrowly one looks. You and I may well have the same expected reliability writ large but very different expected reliabilities in particular domains. I may have better vision; you may have a better sense of smell. I may be better at geometry; you may be better at algebra. I may be sharper in the mornings; you may be sharper in the evenings. And on and on. This is what we might call the Generality Problem for peer disagreement.

Reliabilist theories of knowledge are famously prone to such trouble.⁴ A mathematician can perfectly well have a reliable belief-forming mechanism in her domain of expertise, and so have mathematical knowledge, but be hopeless when it comes to politics. Her overall reliability across both domains is not the relevant measure—that would be too general a measurement. But it is equally problematic to focus too narrowly on the specific case in question. Suppose Tommy is for the most part hopeless at mathematics but that he happens to get Problem #9 right—but only because it is a word problem, and because Tommy always chooses answer *c* when given word problems, and because as it happens the answer

to cognitive reliability. More broadly, one might also include possessing the same information. Sameness of information, or sameness of evidence, is itself a problematic feature of the Impartiality thesis, but for purposes of this paper I am largely setting it aside.

³Here I follow Roger White, "On Treating Oneself and Others as Thermometers," *Episteme*, vi, 3 (October 2009): 233–50, at pp. 235–36.

⁴See Earl Conee and Richard Feldman, "The Generality Problem for Reliabilism," *Philosophical Studies*, LXXXIX, 1 (January 1998): 1–29.

to #9 is *c*. In cases of this exact kind Tommy is extremely reliable, because he reliably chooses *c* in such cases and because *c* is indeed the answer in all cases of *this exact kind*. Obviously, we are individuating cases too narrowly for a reliabilist theory of knowledge to be workable. No one would suppose that Tommy knows the answer to #9. But it is surprisingly hard to see what the correct level of generality is.

It is similarly unclear, at first glance, how to measure reliability in cases of peer disagreement. Descriptions of what it is to be an epistemic peer tend to restrict the scope of reliability to a particular domain but have paid little systematic attention to just what level of generality is called for. Obviously, some restriction is appropriate. If the disagreement concerns Brazilian politics, what matters is whether you and I are epistemic peers in that domain, not whether I am more reliable in the domain of baseball. But just how narrow should we go? Is the relevant measure of peerhood our reliabilities with regard to politics in general, or South American politics, or Brazilian politics? I hope to show that these questions admit of a fairly precise answer. Whereas a reliabilist theory of knowledge is under pressure to be neither too general nor too specific, assessments of disagreement should measure reliability as finely as possible, taking into account all of the relevant factors, and tailoring the assessment of reliability narrowly to the particular matter in dispute. If we disagree about whether it will snow here tomorrow, then we might begin by considering our respective overall reliabilities regarding weather forecasting. But if we discover that one of us is more reliable about *next-day* forecasts, or more reliable about *snow*, or more reliable about the weather *here*, then it is this information that matters. In general, what we want is the most narrowly tailored information available about one's expertise regarding the particular case in question.

The significance of this conclusion—as well as its correctness—will emerge in the discussion to follow. But, as will also become clear, measuring reliability in the context of disagreement is not quite as straightforward as these initial remarks suggest, and it seems only fair to acknowledge some of the difficulties right away. One complication is that there will tend to be limits on how specifically we can gauge an agent's reliability. If you and I disagree about the name of the current president of Brazil, then we would ideally like to know which of us is more reliable when it comes to this very topic, names of Brazilian presidents. But we may have to settle for a comparison of our more general reliabilities about world politics. How do we know exactly what we have to settle for? One key here

is to remember that the relevant issue is our *expected* reliability. Since our overarching concern is to find the rational response to disagreement given the rest of what an agent believes, the relevant question for me to ask myself about reliability is this: what degree of reliability do I have reason to take myself to have here, and what degree do I have reason to take you to have? Although we would ideally like the most fine-grained assessment possible, I am likely to have information only of a more coarse-grained kind, based perhaps on memories of conversations with you about assorted matters of world politics. Such contingencies of available information complicate matters but also permit an approach to the generality problem that is not available to reliabilist theories of knowledge. Because the reliabilist cares only about the external fact of reliability, there is no downward constraint on specificity, until we arrive at the limiting, uninformative case of reliability in this single instance. In contrast, because peer disagreement focuses on expected reliability, there will be downward constraints on specificity in any real-world case.⁵

A second complication concerns an essential qualification to the rule that we should consider reliability in the most fine-grained way available. If you and I disagree about the name of the Brazilian president, then the most fine-grained way for me to evaluate your reliability is in terms of how likely I take you to be right about this particular question. There, however, I assess your expected reliability as low, inasmuch as I think your answer is wrong. Obviously, though, that is too specific an assessment for present purposes, since it would blatantly beg the very question at issue—how much weight to give the contrary opinions of one's peers. The solution must be somehow to measure reliability while setting aside that which is disputed. But it is not obvious how to maintain a sufficiently fine-grained approach when one leaves out of account everything the two sides disagree on. Here lies the heart of my concerns, and here is where it will be useful, in section IV, to move behind the veil of ignorance.

⁵ If, however, we somehow could form reasonable expectations about an agent's reliability in some hyper-specific scenario, then it would be rational to follow those expectations in adjusting our own beliefs. This shows that the difference with reliabilism is not just a matter of being able to resist being pulled down to the limiting case of a single instance, but more fundamentally a matter of the clash between externalist and internalist approaches. It is, so far as I have found, universally accepted in the peer-disagreement literature that the appropriate approach is internalistic—how we ought to adjust our beliefs in the face of disagreement is a function of what information we have access to. I relegate these remarks to a footnote because they raise issues I mean to take for granted in what follows.

II. NEGLECTED EVIDENCE

Solving the Generality Problem for peer disagreement allows us to solve two of the most prominent objections to the Impartiality thesis. I consider first the objection from neglected evidence and then, in the following section, the objection from absurd disagreements.

The neglected-evidence objection begins by observing that, if we both possess the same information and yet disagree, then Impartiality demands that our disagreement get decided by weighing my credence against yours, and my reliability against yours. If these are equal and opposite, then we each should arrive at a credence of 0.5. What looks objectionable is that we would seemingly have arrived at this result without considering the evidence on which we base our respective opinions.⁶

This way of putting the objection is, however, not entirely apt. After all, if my evidence that p really were to drop out of the picture, then I would lose any reason to believe p , and then all I would have to go on is your belief that $\sim p$. Since I take you to be just as trustworthy as me, it would then be rational for me to follow you and believe $\sim p$. We would have a straightforward case of testimony, not disagreement. The reason I go to 0.5, rather than embrace your view, is that I am still paying attention to my evidence. There is, however, a better way to formulate the neglected-evidence objection. For although Impartiality does not ignore the evidence, it does seem to ignore the evidence's strength. If I determine that you and I are locked in genuine peer disagreement, then Impartiality takes me right to 0.5, regardless of how strong or weak my evidence looks to be. Equal credences and equal reliability between peers automatically yields stalemate, no matter what the evidence. That cannot be right.

Roger White has developed this line of objection in some detail, and it will be worthwhile to work through some of the details of his account. On White's analysis, the conditional probability I should consider in a case where you and I are locked in peer disagreement over p is not the probability of p given your contrary

⁶The objection has been set out most forcefully in Thomas Kelly, "Peer Disagreement and Higher-Order Evidence," in Richard Feldman and Ted A. Warfield, eds., *Disagreement* (New York: Oxford, 2010), pp. 111–74. As Kelly puts it, "With respect to playing a role in what is reasonable for us to believe at time t_1 , E [the evidence] gets completely swamped by purely psychological facts about what you and I believe" (p. 124). See too David Enoch, "Not Just a Truthometer: Taking Oneself Seriously (but not Too Seriously) in Cases of Peer Disagreement," *Mind*, cxix, 476 (October 2010): 953–97, at p. 969: "the Equal Weight View requires that in the face of peer disagreement we ignore our first-stage evidence altogether."

belief, but the probability of p given both your contrary belief and the evidence:

$$(1) \quad P(p \mid e \ \& \ \text{you believe } \sim p).$$

This seems to take into account both what it is I want to figure out (the probability of p) and what information I have to go on (the evidence and the fact of your contrary belief). Plugging this formula into Bayes' Theorem reveals that we get the desired equal-weight outcome (that is, that the conditional probability here is $\frac{1}{2}$) if and only if we assume that the probability of p given the evidence is exactly the same as your expected reliability:

$$(2) \quad r_y = P(p \mid e).$$

This is not an obvious result,⁷ but one can see intuitively why it makes sense as follows. First, treat an agent's expected reliability regarding p as the conditional probability of p given that the agent believes p . Now, in the present case we are supposing that you and I are locked in peer disagreement: that is, we are in a situation where your expected reliability is equivalent to my expected reliability. Drawing these threads together with (2), we get

$$(3) \quad r_y = P(p \mid e) = r_m = P(p \mid \text{I believe } p).$$

Proponents of Impartiality, if they are to get the desired equal-weight outcome in cases of peer disagreement, must embrace all the equivalencies in (3). The price of not neglecting the evidence—of including e among the information on which I conditionalize in (1)—is that $P(p \mid e)$ gets drawn into (3), as equivalent both to my expected reliability and to your expected reliability. When and only when we do that can we get a value of $\frac{1}{2}$ for (1).

One striking implication of this result is that the probability of p given the evidence is equivalent to the probability of p given that I believe p . On its face, this seems like no bad thing at all—it looks like my beliefs are simply tracking my evidence. White agrees that, so far, we have arrived at nothing more troubling for Impartiality than what “can seem like a bit of common sense.”⁸ But White thinks that further reflection on the situation reveals serious problems. If we are committed to the above probabilistic connection between belief and evidence, then this, according to White, imposes a general

⁷ For the details of how Bayes' Theorem yields this outcome, see White, “On Treating Oneself,” *op. cit.*, p. 239.

⁸ *Ibid.*

constraint on our credences, even prior to disagreement.⁹ He calls this the Calibration Rule:

If I draw the conclusion that p on the basis of any evidence e , my credence in p should equal my prior expected reliability with respect to p .¹⁰

The proponent of Impartiality is stuck with this result, inasmuch as it is embedded in (3). This is the fruit of our attempting to take count of the evidence, back in (1), but it is here, for White, that the neglected-evidence objection finally becomes vivid. The problem, as remarked earlier, is not that this Calibration Rule calls on us to ignore the evidence, but that we seem unable to respond to differences in the kind of evidence available. White considers a situation where I have very strong evidence for p but my expected reliability for p is only 70%. The Calibration Rule seems to require that I downplay the significance of the evidence and maintain a credence in p of 0.7. The result, as White puts it, is that “the *strength* of the evidence—in this case, the fact that it strongly supports p —has no role to play in determining my attitude.”¹¹ Here White is surely right: this must be wrong.

The situation, then, is that Impartiality requires a policy of proportioning one’s reliability to one’s evidence. But sometimes it seems clear that the evidence warrants greater confidence than one’s expected reliability predicts. To make this vivid, consider disagreement over an arithmetic test. You and I discuss the answers afterwards, and we find that our answers agree except for Problem #9. I regard you as my epistemic peer, so whereas I had been thinking that I aced the test, now I fear that there is a good chance I got a question wrong. But now suppose I still have the question sheet, and I look at #9. It’s a word problem, and I pride myself as being really good at word problems. Moreover, when I look at this particular problem, the answer seems *clear* to me. So now it looks like I should be feeling good about my chances. No doubt I should not entirely ignore our disagreement, but surely I should not drop my credence regarding my answer to #9 down to $\frac{1}{2}$, as Impartiality would seem to require. What I should do instead is to let my evidence increase my credence in my answer above my expected reliability, which will cause me to give less than equal weight to your answer. This violates the Calibration Rule.

⁹ I take for granted the usual direct relationship between probabilities and credences. For the sake of clarity and vividness, in some places I treat belief as admitting of degrees expressed in terms of credences, and in other places I treat belief as all or nothing. At the cost of some awkwardness, one could formulate my conclusions consistently in one way or the other.

¹⁰ White, “On Treating Oneself,” *op. cit.*, p. 239.

¹¹ *Ibid.*, p. 240.

White is quite right that Impartiality demands adhering to the Calibration Rule. But the correct moral to draw here is not that Impartiality is in trouble, but that we need to distinguish between more and less fine-grained measures of an agent's reliability—that we need to grapple with the Generality Problem. If I am brilliant at word problems and terrible at graphing, then my overall expected arithmetical reliability may be 90%, and that may be a useful number to know, but it may also be useful to have a more fine-grained measure of expected reliability, which shows, for instance, that I am 60% reliable on graphing problems and 98% reliable on word problems. The Calibration Rule is defensible provided we consider sufficiently fine-grained measures of reliability. Before taking the test, it is reasonable to have a credence of 0.9 in my answer to the first problem. Once I see that it is a graphing problem, however, my heart ought to sink, and I should revise downward. How far? To 0.6, of course, no more and no less. But what about once I work through the problem? Since I am bad at graphing, this may make me even more dispirited, lowering my credence in p still further. But if this is how I always feel when I do graphing problems, then it would be irrational to become less confident—I should stay at 0.6. Conversely, suppose that working through the problem gives me a sense of confidence in my answer. If *this* is how I always feel, then that confidence too should count for nothing. I should stay at 0.6. Of course, we might need to develop even more fine-grained measures of reliability to distinguish between the various degrees of confidence I might feel in light of the evidence. If one's measure of reliability is not nuanced enough to distinguish between those cases, then of course looking at the evidence will make a difference to one's credences. Without localized information, I will often violate the Calibration Rule, adjusting my credences upward and downward, from case to case, but in a way that will cohere in the long run with my antecedently predicted global reliability. If I manage to arrive at sufficiently localized information about my expected reliability, then I should conform my credences to the Calibration Rule.

What does this show us about peer disagreement, and the neglected-evidence objection? Return to our disagreement over Problem #9. If this is a case of peer disagreement, then *ex hypothesi* our expected arithmetical reliability is the same. But suppose I now look at my question sheet and see that #9 is a word problem. My credence in my answer goes up, and I have violated the Calibration Rule relative to that initial, rough-grained measure of my reliability. But then I remember our disagreement. The question I need to ask, of course, is how good you are at word problems—I need to make a

more fine-grained assessment of *your* expected reliability. If that assessment leads me to expect that we are equally reliable at word problems, then I should return to giving our answers equal weight. It is as if the evidence has dropped back out of the story. If, instead, I conclude that I am better at word problems than you are, then I should not give your view equal weight, but then we would no longer be locked in peer disagreement. In this localized context, I would have concluded that you are not my peer. Either way, the evidence does not really drop out but instead gets assimilated into other measures. If we stipulate that you and I are equally reliable agents who share all the same information, and we stipulate that we are assessing reliability in a sufficiently fine-grained way, then disagreement between us rationally requires giving our views equal weight. The character of the evidence plays a role twice here, both in explaining our disagreement and in calibrating our fine-grained reliabilities.

This solution to the neglected-evidence objection shows why the Generality Problem for peer disagreement must be solved by using fine-grained estimates of expected reliability. The solution in fact sets a minimal condition on fine-grainedness: expected reliability must be measured finely enough as to allow us to adhere to the Calibration Rule. My judgment that the two of us are peers, in terms of our expected reliability, should lead to a credence of 0.5 in cases of disagreement over p if and only if my reflecting on the evidence does not lead me to a credence in p that diverges from my expected reliability. Where it does, I need to recalibrate whether we are peers by attempting a more localized estimate of reliability—of both mine and yours. Only once I have reached a sufficiently fine-grained assessment of whether you and I are peers in this particular kind of situation, given this particular sort of evidence, do I have good reason to give your view equal weight.

The demand for fine-grainedness may look on its face as if it solves the neglected-evidence objection by making cases of true peer disagreement exceedingly hard to find. Your reliability may be the same as mine in arithmetic overall, but that counts for nothing, as far as assessing disagreement goes, unless our reliabilities correspond all the way down to the most fine-grained level, so that I have to consider whether your reliability matches mine on this very problem, on this very day, at this very time of the morning. What was supposed to be a pervasive phenomenon may look like quite a special, *recherché* case. Here again, however, it is crucial to remember that what matters is my and your *expected* reliability, from my (and your) very limited vantage point. For peer disagreement to disappear in the way just imagined, we would have to be able to make comparative

assessments of reliability in an exceedingly fine-grained way. Perhaps this is how it is for the angels—perhaps they understand their own minds and the minds of their fellow angels so well as to leave no room for doubt about the true celestial hierarchy of intelligences. Perhaps, as a result, the angels are able to spend more time teaching and learning and loving, and less time arguing. For us, the situation is sadly otherwise. Although I can discern myself to be better at some kinds of math problems than others, and a better philosopher in some areas than others, I am frequently unable to discern whether I am comparatively better at such things than you are. Given such limited insight into questions of comparative reliability, it is often reasonable to treat one another as epistemic peers.

III. ABSURD DISAGREEMENT

A second prominent line of objection to Impartiality arises from cases of absurd disagreement—that is, situations where the view of one’s opponent seems not just false but manifestly false. Consider this case, modeled after an example from Jennifer Lackey:

Maureen and Tim disagree about where the restaurant My Thai is located. Maureen insists that it is on Michigan Avenue, and vividly remembers walking past it there on countless occasions, and eating there many times. Tim is equally adamant that it is on State Street, and appeals to equally vivid memories. Each has been living in Chicago for many years and knows the city well. Discussion makes it clear that neither is supposing the restaurant to have recently moved; both are contending it has been in that location for a long time; both seem to be entirely serious. They are clearly talking about the same streets and the same restaurant.¹²

From Maureen’s perspective, this disagreement is absurd because she has utterly vivid memories of My Thai’s location on Michigan Avenue. She is nearly as confident about the restaurant’s location as she is about where she herself lives. Lackey takes this to be a case where Impartiality goes wrong. Given how certain Maureen is about the truth of her position, it would not be rational to give Tim’s view equal weight, even though she would have antecedently accepted him as her epistemic peer with respect to local geography. Lackey concludes that Impartiality fails in cases like this, because Maureen’s confidence that she is right ought to overcome her antecedent belief that she and Tim are epistemic peers in this domain.

¹² This is my version of the case described in Lackey, “A Justificationist View,” *op. cit.*, p. 308.

Lackey gets this result because she too is working with the wrong resolution of the Generality Problem. She supposes that Impartiality assesses whether the two agents are peers in general on topics of this kind, independently of the present disagreement. As she puts it, the two agents count as peers “in the abstract,” even though “in the context of the actual disagreement itself” this judgment of peerhood collapses.¹³ When the situation is so conceived, then of course Impartiality will fail. After all, prior judgments of expected reliability get overridden all the time by disagreement. It is perfectly common to suppose someone one’s epistemic peer until he opens his mouth, and then consequently to ignore his views entirely. But this is not how the Impartiality thesis should be spelled out. One needs to take into account as much information as possible, including the present disagreement and all of its attendant circumstances, and then make an overall decision about what is rational. It is clearly appropriate for Maureen to weigh her confidence in her memory against her confidence in Tim’s knowledge of Chicago, and against Tim’s own memories, and she might well be right to downgrade Tim’s reliability as a result. This is what Lackey says *she* would do in such a situation, and that seems both plausible and consistent with Impartiality, properly understood, since Impartiality requires assigning equal weight to another’s view only if the other is one’s epistemic peer *in the present situation*.

It is also imaginable that Maureen, on reflection, would start to doubt herself. If Tim, in discussion, shows every sign of being equally confident and competent, then Maureen might reasonably come to fear that she’s as likely to have gone badly astray as Tim is. Lackey suggests that this is the less rational response, on the grounds that Maureen will have personal information about her own case that she will not have about Tim: details about her own evidence that she cannot articulate, for instance, and knowledge about her own current

¹³ *Ibid.*, p. 314. Lackey is quite explicit in supposing that the Impartiality thesis must measure an agent’s reliability generally, prior to the disagreement. She treats it as a constraint on giving equal weight that “the disagreement itself should not change one’s beliefs about the probability that one is right” (*ibid.*, p. 313). In support of this reading, she quotes Elga, “Reflection and Disagreement,” *op. cit.*, p. 488: “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 you think it that you are correct? The equal weight view says: 50%.” Lackey does not notice that, later in that same paper, Elga proceeds quite differently, making the outcome “conditional on what you later learn about the circumstances of the disagreement” (*ibid.*, p. 491). He in fact employs this move precisely to handle the sort of case where “an advisor you treated as a peer comes up with a conclusion that you find utterly insane” (*ibid.*, p. 490). I consider the details of Elga’s approach in section iv.

mental acuity that she will not have about him.¹⁴ In some kinds of cases such information may well be decisive. In this restaurant example, for instance, it seems particularly likely that Tim is either joking or not really paying attention, possibilities that Maureen can categorically dismiss in her own case. It seems doubtful, however, that this sort of first-person information creates such an asymmetry across the board, since we face equally serious handicaps in justly appraising our own reliability. If Tim, for instance, were starting to lose his mind, Maureen would likely have heard people talking about it. But if Maureen herself were losing her mind, she might well be the last to know, or might never know. Such self-blindness is particularly poignant in cases of mental deterioration, but extends more widely. We notoriously exaggerate our own competences. And the difficulty in getting honest feedback about oneself makes it, in many contexts, easier to evaluate the reliability of others than to evaluate one's own reliability.¹⁵

Impartiality can welcome all of these reflections and can accept any number of outcomes in the restaurant case. Both Maureen and Tim may rationally continue to insist they are correct, or one or both may decide to abandon their belief in light of the disagreement. All that Impartiality insists on is equal consideration for the views of others, which is consistent with giving one's own view more (or perhaps less) weight when the circumstances warrant it. It is only if one of the two regards the other as an epistemic peer that she or he must weigh the other's view equally to her or his own. Lackey gets a violation of that principle because she assesses peerhood at a general, abstract level, independently of the disagreement itself, but then assesses what ought to be believed in light of that disagreement. This is essentially the same mistake made by the neglected-evidence objection; indeed, one might think of these as alternative versions of the same objection. Each employs the strategy of measuring reliability in a general way and then introducing specific details that make Impartiality look implausible. The solution in each case is to insist on a sufficiently fine-grained measure of reliability that takes into account the specific circumstances of the disagreement.

¹⁴ "How often does it happen, for instance, that I know that my colleague...is not depressed, exhausted, distracted, and so on, on any given day?" (Lackey, "A Justificationist View," *op. cit.*, p. 311). Rarely, she thinks. For similar remarks see David Christensen, "Disagreement, Question-Begging and Epistemic Self-Criticism," *Philosophers' Imprint*, xi, 6 (March 2011): 1–22, at pp. 9–10.

¹⁵ For a useful popular discussion of the various obstacles to self-evaluation, see Thomas Gilovich, *How We Know What Isn't So: The Fallibility of Human Reason in Everyday Life* (New York: Free Press, 1991), chapter 7.

IV. BEHIND THE VEIL OF IGNORANCE

These two objections against Impartiality fail for the same reason: they presuppose the wrong solution to the Generality Problem. When comparative reliability is measured finely enough, in light of the current disagreement, the objections dissolve. It is, however, not perfectly clear what counts as “finely enough” for purposes of measuring expected reliability. We have, in section II, discovered one minimal condition: measures of expected reliability must be sufficiently specific to satisfy White’s Calibration Rule. This leaves unclear, however, what lower limits there might be to the appropriate degree of specificity. We have seen that the specific circumstances of the present disagreement cannot be entirely ignored, but surely some matters must be held in abeyance. If nothing else, I am obviously not entitled to take for granted the truth of p , when p is the thing we disagree about. To suppose that p is true begs the whole question.¹⁶

David Christensen, in an effort to get clear on the proper formulation of Impartiality, offers what he calls the Principle of Independence:

In evaluating the epistemic credentials of another’s expressed belief about P, in order to determine how (or whether) to modify my own belief about P, I should do so in a way that doesn’t rely on the reasoning behind my initial belief about P.¹⁷

Although Christensen supports Impartiality, this principle looks on its face to lead directly to the sorts of objections we have been considering, inasmuch as it would seem to preclude an agent from considering the present disagreement. To say that I cannot “rely on [my] reasoning” seems precisely to enjoin neglecting the evidence, and seems to preclude discounting absurd disagreements. So understood, Christensen’s principle would get the Generality Problem wrong.

But Christensen himself does not understand his Independence Principle in quite this way. The point of the principle, Christensen says, is to block agents from blatantly begging the question about the proposition in question. In saying that I should evaluate my

¹⁶ Although this point might seem uncontroversial, it is surprisingly denied by Enoch, “Not Just a Truthometer,” *op. cit.*, p. 982. In a case where I believe p and you, my epistemic peer, believe $\sim p$, Enoch holds that I should not give your view equal weight, simply for the reason that p is true. He seems not to mind begging the question in this way, on the grounds that otherwise we will be mired in skepticism. I would suggest to the contrary that, if skepticism does follow, then in cases where we find ourselves locked in irresolvable peer disagreement it is better to admit those consequences forthrightly.

¹⁷ Christensen, “Disagreement, Question-Begging and Epistemic Self-Criticism,” *op. cit.*, pp. 1–2.

opponent's reliability without letting myself "rely on [my] reasoning," he seems to mean only that I should not rely on the *conclusion* of my reasoning. I may consider the immediate circumstances of the disagreement, and may even reflect on the path that led me to my conclusion, and ~~to~~ compare that path to my opponent's path. All this information is allowed in calculating my and my opponent's expected reliability. What I may not do is take for granted the correctness of my conclusion.¹⁸ Hence Christensen's Principle of Independence does after all allow assessments of reliability that are sufficiently fine-grained to handle both cases of absurd disagreement and the charge of neglecting the evidence.



Unfortunately, the lower limits of the Generality Problem cannot be defined as simply as this. The Impartiality thesis requires that parties to a dispute set aside more than just their contending beliefs regarding p ; they also must to some extent set aside the reasoning that led them to those beliefs. Suppose I am going over my answer to Problem #9 in an effort to sort out why you and I disagree. I am prepared to set aside my answer, but in going over my reasoning step by step I may well become all the more confident that this answer is correct. Suppose that you go over your reasoning at the same time and become all the more confident that your answer is correct. If we judge ourselves epistemic peers, then Impartiality will recommend that we each adopt a credence in p of 0.5. But for that to seem right to me, I must set aside more than just my antecedent confidence in p . I also have to set aside my confidence in the reasoning that led me to p . For we can imagine that it might be clear as day that, if my reasoning is correct, then p clearly follows, whereas if your reasoning is correct then $\sim p$ follows. Disagreements of all kinds have just this structure, where the two parties disagree about p in virtue of some prior disagreement over a principle that plainly entails p . Hence Impartiality often requires setting aside more than just the claim at issue; it seems also to require setting aside the truth of all the steps in reasoning that led to that claim.

¹⁸ Christensen, "Disagreement, Question-Begging and Epistemic Self-Criticism," *op. cit.*, p. 10: "I relied on the claim that I arrived at my answer to the math problem by a very reliable method. But my reasoning did not rely on the results of my calculations at all. I did not say, 'Well I'm very sure the answer is \$43. My friend says it's \$45, so something screwy must have gone on with her.' That sort of reasoning would indeed violate Independence." For a recent extended argument against Christensen's principle, see Thomas Kelly, "Disagreement and the Burdens of Judgment," in David Christensen and Jennifer Lackey, eds., *The Epistemology of Disagreement: New Essays* (Oxford: University Press, 2013), pp. 31–53.

For an attempt to honor this constraint, we might consider Adam Elga's version of the Impartiality thesis. Elga argues that an assessment of comparative reliability between yourself and a rival must take into account "the circumstances of the disagreement." But he insists that these circumstances "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..."¹⁹ So in considering Problem #9, I am allowed to consider the kind of problem that it is, the extent of my confidence in my answer, and the care that I took in answering the problem. I should consider all of these "circumstances" and compare them with the circumstances of your approach. But I am prohibited from taking into account the steps in my reasoning, because that would lead me to insist on my answer to #9, rather than giving equal weight to your answer. On Elga's approach, "thinking through the disputed issue" effectively begs the question.

But Elga's approach also fails to resolve the low end of the Generality Problem, because it excludes factors that are vital for a rational appraisal of some disagreements. Suppose Anna has no information about Tommy's ability at math. Since she regards her own abilities as average, she reasonably expects him to be her peer. Then Anna learns that they gave different answers to Problem #9, which causes her rightly to fear she got that problem wrong. But suppose Anna now considers "a detailed specification of the chain of reasoning" that led her and Tommy to their respective conclusions. She considers that whereas she tried to solve the word problem by plugging the constants into an equation and then solving for the variable, Tommy's method was simply *if it's a word problem, go with c*. Recognizing this, Anna clearly should cease thinking of Tommy as her peer. But Elga's approach, absurdly, would not let her consider these specific features of the disagreement.

We face a puzzle, then, of the following shape. On the one hand, a rational appraisal of disagreement sometimes requires that I compare my line of reasoning against your line of reasoning. Such a comparison may sometimes give me reason to discount your answer, as in the Anna–Tommy case, even if I had antecedently judged you to be my peer. On the other hand, taking my line of reasoning into account sometimes seems to involve begging the question just as surely as would taking into account the truth of p itself. That

¹⁹Elga, "Reflection and Disagreement," *op. cit.*, p. 490.

line of reasoning, after all, may be precisely what generates our dispute. It is easy to see, then, why Elga would entirely exclude reflection on the details of one's reasoning. And perhaps this is why Christensen's Independence Principle insists rather vaguely that I may not "rely on [my] reasoning."

What we want, it seems, is for parties to a dispute to be able to consider *everything* about the dispute, on down to the most precise details that might be relevant, but somehow without begging the question in favor of the truth of their own views. This would honor the initial idea that the Generality Problem for peer disagreement should be solved in the direction of maximal specificity. But how can I take into account *all* the details of a dispute over p , without thereby assuming the sort of pro- p attitude that is to beg the question? My suggestion is that we borrow a page from political theory and consider cases of peer disagreement from behind a veil of ignorance. In that spirit, let us imagine ourselves informed about all the factual circumstances of the situation—what is agreed upon, what is contested, what the opposing arguments are, what the credences are of the contending parties—but without knowing *who we are* in the dispute. To say that we would not know who we are means more than that we would be blocked from knowing certain autobiographical facts. It means as well that we would know neither where we stand on the proposition in dispute, nor on any other relevant propositions that are contested by the two sides. We would recognize that one of the parties takes p to be highly likely and regards p as supported by strong evidence, but we would also recognize that the other party thinks none of these things. We ourselves would enter imaginatively into a state of neutrality on such questions, setting aside our intuitions in one direction or the other.²⁰

Here then is my ultimate proposal for how to solve the Generality Problem for peer disagreement. In evaluating my reliability in comparison with yours, I should make that evaluation in the context of the very dispute in question, taking into account all the available information about how I made my decision and you made your decision, and my confidence versus your confidence—leaving out of this picture only the self-locating facts about where I am in

²⁰ The sort of information to be excluded could be characterized more precisely in terms of the theory of self-located beliefs set out in John Perry, "Frege on Demonstratives," *Philosophical Review*, LXXXVI, 4 (October 1977): 474–97; and subsequently developed in terms of centered worlds by David Lewis, "Attitudes *De Dicto* and *De Se*," *Philosophical Review*, LXXXVIII, 4 (October 1979): 513–43. Of course, only the *disputed* self-locating facts must be excluded.

the debate. Insisting that disagreement be adjudicated from behind such a veil is a way of honoring the core motivation for Impartiality: the idea that my own beliefs should not get special weight just because they are my own, or just because they follow from what seems right to me. Of course I must ultimately believe the things that strike me as true—there is no other course of action—but Impartiality insists that we should modulate those judgments in light of our disagreements with other seemingly intelligent people. Reasoning from behind the veil allows us to consider enough of the details of those disagreements to see when we should be worried by disagreement and when we should ignore it, while blocking the distorting influence that arises out of self-bias.

Political theorists differ over exactly what information should be allowed behind the veil. It should be clear by now that, in cases of peer disagreement, it is essential to allow full information about the circumstances of the dispute and the nature of the evidence, but also essential to block self-locating information with regard to all the contested issues. So far, I have mentioned as contested only belief in p itself and in those propositions that serve to support or rebut p . There may, however, be more than this that should be excluded. When you and I disagree, I may suspect that the real reason you think you are right is not because your arguments are stronger, but simply because you are white or male, or because your degrees are from the Ivy League, or because you have an impressive job. Although the literature on peer disagreement has hitherto ignored such issues, they clearly play a role in disagreements of all kinds. The veil does not entirely exclude such sociological facts from consideration, inasmuch as they are among the details that are potentially relevant to a full evaluation of the dispute. What the veil excludes are any disputed attitudes toward such facts, when those attitudes are relevant to the disagreement. Applying this heuristic of a veil to the phenomenon of peer disagreement allows us to give a place in our theory to these kinds of social biases.

In political theory, it is sometimes objected that there is something problematic about decision-making behind a veil. Such an objection might arise here too, in three kinds of ways. First, one might want a fuller account of what information gets excluded. No doubt there is much more that might be said about that, but I will not attempt it here. Second, one might complain that it is simply not possible to screen off the sort of self-locating information I propose to exclude. As a psychological matter, no doubt that is true. Barring radical neurological failure, it is not humanly possible, even for a moment,

to enter fully into the mindset that the veil prescribes. We know who we are, and how the evidence strikes us, and we cannot *really* forget that. What the veil describes, then, is an epistemic ideal to which we might aspire in attempting to respond rationally to disagreement.²¹ Third, one might charge that the veil, beyond describing an unachievable ideal, offers a heuristic that cannot coherently be followed in any way. Here I must disagree. Given that we cannot *really* forget who we are and what we think, the heuristic of the veil asks us, in effect, to engage in reasoning for a time without adopting a particular perspective. This seems no more mysterious or problematic than the familiar request that we, for the purposes of an argument, *suppose that p*, and see what follows. Instead of making a supposition, the veil asks us to refrain from certain suppositions, and to see what follows from that. This will not necessarily be easy to do, inasmuch as it will often be difficult to figure out just what self-locating information needs to be walled off, and difficult to know whether one *is* in fact walling it off. But the fact of such difficulties is no objection to the heuristic. Indeed, our chronic tendency to respond irrationally to disagreement should lead us to expect that the proper approach often will be difficult to follow.

To put these remarks on a more concrete footing, let us revisit some cases. If Maureen puts herself behind the veil, she will be able to consider all the facts about her confidence in the restaurant's location, just without taking into account that *she* is the one who is confident. If she does not know much about Tim's evidence and confidence level, then even from behind the veil she would rationally adhere to her own belief. But if she had information that Tim's belief was based on similarly strong evidence and held with equal confidence, then she should judge herself uncertain about the correct answer, having no basis for choosing between one person's confidence and another's. The cases being symmetrical, from behind the veil, she would have no grounds to favor one side or the other. That intuitively seems rational.

Anna, from behind the veil, sees two methods for solving the word problem: the familiar approach taught in the schools, and Tommy's idiosyncratic method. Clearly my proposal would be in some trouble if Anna, behind the veil, were forced to give Tommy's answer equal weight. We would then seem to be confronted once again with the absurd-disagreement objection against Impartiality. In the real world,

²¹ On the place of ideals in epistemology, see Robert Pasnau, "Epistemology Idealized," *Mind*, cxxii, 488 (October 2013): 987–1021.

however, there is no risk of that outcome. Anna, in imagining herself behind the veil, has to adopt a stance of provisional neutrality regarding any relevant matters of disagreement between her and Tommy. But if we are talking about the world as we know it, then we should not suppose that Tommy is going to adhere steadfastly to his method. He employs his method, we can assume, for lack of anything better, and we can assume that he would cheerfully concede Anna's method to be vastly preferable. (Indeed, we might think that Tommy never even *believed* his answer.) Hence in any real-world case of this kind, the comparative reliability of the two methods would not be a contested issue, and so there would be no need for Anna to remain neutral on the two methods. Anna could embrace her own method from behind the veil, and Tommy too could embrace Anna's method, and both could come to share Anna's original credence in her answer.

The hard case is the otherworldly situation where Tommy stubbornly believes that his method is reliable. In that case it can look as if Anna must, behind the veil, give the two methods equal weight, inasmuch as she must set aside her own views about which method is preferable. But for this to stick as an objection against Impartiality, we would have to formulate the example in even more otherworldly terms. For in the world as we know it, there will be other agents beyond Tommy whose views Anna can factor into her thinking, and whose preference for her orthodox textbook approach will swamp whatever weight Tommy's idiosyncratic method might carry. There will also be a vast number of background beliefs that make his method seem incredible, and even if these background beliefs are themselves controversial between her and Tommy, there will again be countless agents on Anna's side, whose agreement with her will swamp whatever weight her disagreement with Tommy might have. For the Anna–Tommy case to cause trouble, then, we would have to imagine Anna's world being such that she has access to the views of no one other than herself and Tommy, or else we would have to imagine Anna's living in a world populated by a great many people who think like Tommy do, and share all the bizarre background beliefs required to make his method coherent. These worlds are so different from our own that it is hard to know what to say about them. But if we had to formulate a principle for reasoning under such circumstances, it is at least credible that an agent such as Anna should suspend belief in her answer to Problem #9, along with a great many other beliefs. In such a world, Anna might well not know what to think about much of anything. Even in such an otherworldly context, then, Impartiality arguably gives the correct answer.

V. LIFE VERSUS TRUTH

Applying the veil of ignorance to the present context illuminates the very different ways in which disagreement gets handled in political theory and epistemology. I have argued that the veil helps us see how Impartiality yields the rational response to disagreement. But of course Impartiality, in many cases, yields the conclusion that we should *suspend* belief. This is a curious result in the present context, since political theorists employ the veil of ignorance to achieve *consensus* over a particular political arrangement. Rawls, for instance, supposes that individuals behind the veil will agree on his two principles of justice; others have fashioned veils that promote utilitarianism, and so on. Such accounts obtain consensus because of what they do and do not allow behind the veil, and because of how they require parties behind the veil to reason. Rawls's "thick" veil excludes not just the place of individuals in society, and their own moral and political views, but also excludes information about the character of the moral and political disagreement itself. Individuals behind the veil are supposed to make their decision based only on their rational self-interest, albeit from their veiled perspective. Under these tightly restricted conditions, it is plausible to suppose that rational agents would agree on a theory of justice.²²

The epistemologist gets different results by conceiving of the situation in a fundamentally different way. In the context of peer disagreement, the goal is to proportion one's credence in p to the probability of p given the evidence. The veil works as a heuristic device allowing us to set aside factors that are irrelevant and distracting to an accurate consideration of the evidence. Rawls, in contrast, does not seek any such thing. He begins by setting aside a great deal of information that clearly is evidentially relevant, and then requires individuals behind the veil to ask themselves not which theory the remaining evidence best supports but which theory will best serve their self-interest. There is no reason at all to think that conclusions produced under these conditions will be true, or even well-proportioned to the evidence.²³ This sounds like

²²John Rawls, *A Theory of Justice* (Cambridge: Harvard, 1971), chapter 3. For a useful summary, see Samuel Freeman, *Rawls* (New York: Routledge, 2007), chapter 4. For some recent reflections on how to cope with disagreement that persists even behind Rawls's veil, see Ryan Muldoon et al., "Disagreement behind the Veil of Ignorance," *Philosophical Studies*, CLXX, 3 (September 2014): 377–94.

²³See Chandran Kukathas and Philip Pettit, *Rawls: A Theory of Justice and its Critics* (Stanford: University Press, 1990), p. 72: "The point of the contractarian enterprise is not to identify what justice requires....The point is to identify a way of organizing society...which fits with the constraints of the concept of right, in particular the

a damning criticism, but it is not so intended. Rawls's aims are practical, not theoretical, in the sense that he seeks a decision procedure for organizing society. His later writings frankly concede that he does not aim at a *true* theory but only at one that is workable.²⁴ In contrast, the literature on disagreement assumes from the start that our only concern is with fidelity to the truth. It is these strict conditions of epistemic rationality that yield the Impartiality thesis and its accompanying demand to suspend belief in cases of peer disagreement.

Can these two approaches be melded? Epistemology might aspire toward a general theory of disagreement, treating political disagreement simply as a special case and wielding the veil of ignorance as a general tool applied across all domains. In fact, however, as things presently stand, such an aspiration would be in vain, because both the political theorist and the epistemologist pursue narrow, non-overlapping agendas. Political theorists, on the one hand, pursue a workable political consensus. Although this is no doubt a fine thing, comparison with the aims of epistemology reveals its limitations: the veil of ignorance, however it is spelled out, cannot yield a consensus that is both substantive and epistemically rational. Full and impartial rational reflection on the evidence, aimed at maximizing true belief and minimizing false belief, would yield only the consensus that we should suspend belief. Epistemologists, by comparison, are equally limited in their perspective. Their debates takes for granted that the only goal is epistemic rationality: proportioning beliefs to the evidence. This is fine as a theoretical question for philosophers in their armchairs, but in the real world more is at stake than simply achieving the highest possible ratio of true to false beliefs.²⁵

Whether or not we should suspend belief from the perspective of ideal epistemic theory, it is necessary as a practical matter that we act.

publicity constraint." More harshly, see the characterization in Derek Parfit, *On What Matters* (New York: Oxford, 2011), vol. I, p. 357: "Though Rawls's veil of ignorance ensures impartiality, it does that crudely, like frontal lobotomy. The disagreements between different people are not resolved, but suppressed."

²⁴ See John Rawls, "Justice as Fairness: Political not Metaphysical," *Philosophy and Public Affairs*, xiv, 3 (Summer 1985): 223–51, at p. 230: "the aim of justice as fairness as a political conception is practical, and not metaphysical or epistemological. That is, it presents itself not as a conception of justice that is true, but one that can serve as a basis of informed and willing political agreement between citizens viewed as free and equal persons....Philosophy as the search for truth about an independent metaphysical and moral order cannot, I believe, provide a workable and shared basis for a political conception of justice in a democratic society."

²⁵ For further thoughts in this same vein, see Robert Pasnau, "Disagreement and the Value of Self-Trust," *Philosophical Studies* (forthcoming).

Moreover, the point is not just that we each individually need to decide how we will act in the world; we also need to seek consensus, so that we can build institutions together for the common good. So even if abstract theoretical inquiry leads to an impasse, and even if the veil of ignorance itself underwrites that conclusion, there is still a need for some other method of handling disagreement in the real world. The political domain is just one arena where we cannot afford to let our beliefs be governed by the strict demands of epistemic rationality. We may count on the epistemologists, then, to tell us what is most likely to be true. But we need to turn elsewhere for the sorts of conclusions we can actually live with.

ROBERT PASNAU

University of Colorado, Boulder