

Is evidence about one's own doxastic states inert?

(Draft)

Maria Lasonen-Aarnio

Abstract

A lot of epistemologists have recently argued that evidence about the doxastic states of another subject who shares my evidence can give me both evidence about my evidence, and evidence about first-order matters that the evidence bears on. So, for instance, learning that a friend who shares my evidence is very confident that p can give me evidence that my evidence supports p , and evidence that p is true. But assuming that my own states are not perfectly luminous to me, could learning what I think about a matter have the same kind of evidential import? For instance, could learning that I am confident that p give me more evidence about whether p ? It is tempting to say “no”: evidence about my own doxastic states is inert in a way that evidence about the states of others is not. I argue that this is wrong: evidence about my own states is not inert in the way that many think it is. Asking what I think about a matter can be a perfectly legitimate way of gaining more evidence about it.

I Finding out about your evidence

Consider evidence about evidence, evidence that bears on the question of what some (other) body of evidence supports. Few would deny that if I simply don't know what Faye's evidence is, but I do know that she is very good at proportioning her beliefs to the evidence and that she is very confident in a proposition p , this gives me evidence that her body of evidence, whatever it is, supports p . And it may also give me evidence that p . But more interesting are cases in which I have (and know that I have) the very body of evidence on which Faye's opinion is based: in such cases, does the evidence screen out the evidential import that learning about Faye's opinion might otherwise have?

The idea that such screening out always takes place has not been a very popular view of late – at least not when the “higher-order” evidence is evidence about the opinions of one's peers. Consider an analogy. Madame Babineaux utters a sequence of words. Even when I hear her, this need not render evidentially irrelevant what others testify about what she is saying. And indeed, especially if my French is, well, a bit rusty, it is advisable to listen to what others think she just said. Similarly, our access to our own evidence is often limited. One can try to listen to the testimony of the original evidence. But at least sometimes learning what others make of the evidence can provide a further epistemic window to just what its testimony is. For the purposes of this paper I will assume that this is correct.¹

¹ It should be clear that I am assuming there to be objective facts about evidential support. However, nothing I say essentially relies on there always being a uniquely rational opinion that reflects these facts – at least not as long as we can make sense of the idea that learning of the opinions of others can give evidence about which opinions are rational without assuming such uniqueness.

But now assume that my own opinions are not perfectly luminous to me: it is not always rational to be certain that I hold the opinions I do. That is, there is room for genuinely learning about my own opinions. Note that even subjects who are infallible about their own opinions might fail to satisfy luminosity, so long as it is not rational for them to be certain of their own infallibility.² In effect, hardly anyone these days would defend perfect luminosity for actual subjects. It is true that some still think that there is an interesting notion of rationality that requires such perfect luminosity. I heartily disagree, but won't need to settle the issue. I will simply assume that the mere fact that we lack perfect luminosity does not put us so beyond the pale of rationality as to make it uninteresting to ask how we ought to respond to evidence about our own doxastic states.

So here is the question I want to ask: if I can learn about my *own* opinions, can such learning have the sort of evidential impact that, as many authors have recently assumed, learning about the opinions of others can? For instance, can learning that I am very confident that Obama will win the next presidential race give me evidence that the evidence on which my confidence is based strongly supports Obama winning? Moreover, can it give me evidence that Obama will, indeed, win? Many feel an immediate urge to answer these question – especially the latter one – with an emphatic “no”: at least setting aside cases involving some abnormal causal relationship between how things are out in the world and my mental states, evidence about my present opinions is evidentially inert in a way that evidence about the opinions of others is not. It just seems absurd that I could ask myself what I think, and boost my confidence in a claim further as a result of learning about my own opinion. In this spirit, David Christensen (2011), for instance, writes:

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

Those who think that evidence about my present opinions is inert in a way in which evidence about the opinions of others is not may be willing to concede that the opinions of my *past* self (assuming that those opinions are suitably independent of my present ones) can sometimes function like the opinions of other subjects.³ And maybe, if I have completely lost a body of evidence (perhaps due to amnesia) and all I now know is what I believe based on it, my belief can give me evidence about evidence I no longer have. But, the thought goes, when I *have* a body of

² Cf. Christensen (2007).

³ We might imagine the following case: I know that in the distant past, before suffering from amnesia, I have carefully evaluated a body of evidence I now happen to have. I know that I was an equally good evaluator of such evidence as I am now. I then learn that the opinions of my past and present self differ. The case is strengthened by assuming that I know that my past self is, for all practical purposes, exactly like one of my present peers when it comes to evaluating evidence.

evidence – at least setting aside abnormal cases – any such epistemic import of evidence about one’s own states is screened out. If Madame Babineaux is shouting in my ear, asking what *I* make of her speech cannot provide me with further clues as to what she is saying!

This initial reaction may seem to be further confirmed by reflecting on some of the odd-seeming consequences of allowing information about my own doxastic states to have the same kind of import as evidence about the states of others is taken by many to have. Wouldn’t boosting my confidence in a proposition as a result of learning what *I* think involve some illegitimate recycling of my evidence? Couldn’t I keep repeating the procedure, becoming more and more rationally confident in a proposition? And wouldn’t one end up with violations of a plausible synchronic version of the Principle of Reflection?

The purpose of this paper will be to argue that it is at best implausible and at worst incoherent to allow evidence about the (present) doxastic states of others to have a certain kind of epistemic import, but not to allow evidence about one’s own (present) states to have that kind of import. Though I will focus on doxastic states, I very much doubt that the kind of symmetry I will argue for is restricted to learning about such states. Assume that learning that a friend feels moral revulsion towards a certain action can give me evidence that it is wrong; that learning that she fears the next pitch we are about to climb can give me evidence that it is dangerous; and even that learning that she enjoys *A Love Supreme* can give me evidence that it is an excellent record. Then, the kinds of arguments I will give can be expected to generalise to these other cases as well. So, for instance, learning that *I* feel revulsion toward an action could give me evidence that it is morally wrong.

I will first formulate two theses, *First-order inertness* and *Higher-order inertness*, that are attempts to capture the ideas that evidence about my own opinion about whether *p* is inert when it comes to the first-order question of whether *p*, as well as higher-order questions about what my evidence supports. As it turns out, there are strong arguments against these theses, many of which rely on assumptions that numerous epistemologists have recently accepted. After laying out the initial case against the theses, I consider a refined view about the evidential asymmetry between evidence about one’s own doxastic states and evidence about the doxastic states of others. Roughly, the idea will be that if I am rational, and have no reason to think that I am prone to either over- or under-estimate the import of my evidence, then at least *First-order inertness* will hold. I argue that even this refined thesis of first-order inertness is false. Before concluding, I consider and answer some remaining qualms and objections.

II The inertness theses

Let me begin by clarifying what I mean when I talk about learning about someone’s present opinions. Assume that it is now time *t*, and my peer and I hold opinions about whether Obama will win the presidential race, opinions that are based on evaluating a common body of evidence. At a slightly later time *t'* we disclose the opinions we held at *t*. Of course, this gives us new evidence, and in so far as the new

evidence is not inert, it may no longer be rational for us to hold the opinions we did. Hence, if we are rational, the opinions we disclose might no longer be opinions we hold at the time of disclosure. Similarly, assume that at a time t I am wondering how confident I am that Obama will win. Learning about my present opinion will be learning what my opinion is at t . But of course, by the time the learning takes place (at a later time t'), t will no longer be the present moment.

As indicated above, I am assuming that learning about the (present) opinions of others can have both *higher-order* and *first-order* import. For instance, if I discover that my peer believes, based on a common body of evidence, that Obama will win the presidential race, this can give me both evidence about what our common body of evidence supports, and evidence that Obama will win the race. What I want to argue is that if this is right, then the same is true of learning of one's own present opinions. First-person evidence is neither higher-order nor first-order inert.

Assuming that updating happens by conditionalisation, a first pass at the ideas that evidence about one's own states is both higher-order and first-order inert can be formulated as follows. Here Cre_t is the credence function of a (rational) subject at a time t , and P_o is the (or a) rational credence function, a credence function that reflects objective facts about evidential support. I will follow van Fraassen's convention of using lower-case letters for non-rigid designators: " $\text{cre}_t(p) = r$ " should be read as "my credence in p at t is r ", and " $\text{p}_o(p) = r$ " as "the objective, rational credence in p at t is r " (or perhaps, if one doesn't think that there is always a unique rational credence, as "one of the rationally permitted credences in p at t is r "):

First-order inertness

$$\text{Cre}_t(p \mid \text{cre}_t(p) = r) = \text{Cre}_t(p)$$

Higher-order inertness

$$\text{Cre}_t(\text{p}_o(p) = r' \mid \text{cre}_t(p) = r) = \text{Cre}_t(\text{p}_o(p) = r')$$

Unfortunately, these theses run into immediate problems.

Consider *First-order inertness*, and let c be as follows:

c : I assign a credence of 0.8 to some proposition at t .⁴

Assume that, not having perfect access to my credences, at t my credence in c is 0.9. But conditional on my credence in c being exactly 0.8 at t , my credence in c ought to be 1. Hence, $\text{Cre}_t(c) = 0.9$, but $\text{Cre}_t(c \mid \text{cre}_t(c) = 0.8) = 1$. Hence, if we let c be *any* proposition whatsoever, it isn't too difficult to generate counterexamples to *First-*

⁴ In so far as there are uncountably many precise credences I could have, it may be that I ought to assign a credence of 0 to the proposition that I assign a credence of *exactly* 0.8 to some proposition. To avoid such issues, we can assume that talk of assigning a credence of 0.8 to some proposition should be understood as assigning a credence that is within some non-zero interval containing values both below and above 0.8.

order inertness.⁵ A restriction must be placed on the propositions that the thesis is to apply to.

But it is dubious whether restricting the propositions that the inertness theses are to apply to will quite do it. Assume that I am told by highly reputable sources that neuroscientists have ensured, possibly by manipulating my credences, that the following is true: my credence in a proposition is high just in case the proposition is true, and low just in case the proposition is false. Even those who defend the idea that evidence about one's own states is inert in a way that evidence about the states of others is not might well admit that such cases are exceptions, for I have evidence that there is an abnormal causal dependence between whether p and my credence in p . The same kind of point can be made to apply to *Higher-order inertness*: assume that I am told by highly reputable sources that neuroscientists have ensured that my credence in p is 0.9 just in case my evidence supports p to degree 0.5. Then, if I have enough reason to trust the testimony, wouldn't learning that I am 0.9 confident in p give me evidence that my evidence only supports p to degree 0.5?

A further need for refinement might arise from noting that *First-order inertness* is incompatible with a version of van Fraassen's *Principle of Reflection* that applies to one's current credences. Yet, such a principle might seem to some like an uncontroversial part of *Reflection*⁶, capable of avoiding the kinds of counterexamples that a more general version faces:

Current Reflection

$$\text{Cre}_t(p \mid \text{cre}_t(p) = r) = r$$

Note that such a principle faces the same kinds of counterexamples as an unrestricted version of *First-order inertness*. Hence it, too, would have to be restricted. But to see that the two principles are incompatible, assume, for instance, that $\text{Cre}_t(p) = 0.8$. If the subject doesn't know what her credence is, then for some $r_i \neq 0.8$, she will assign a non-zero credence to $\text{Cre}_t(p) = r_i$. Then, the subject cannot satisfy both *Current Reflection* and *First-person inertness*: if $\text{Cre}_t(p \mid \text{cre}_t(p) = r_i) = r_i$, she will satisfy the former but violate the latter; if $\text{Cre}_t(p \mid \text{cre}_t(p) = r_i) = 0.8$, she will satisfy the latter but violate the former. Nevertheless, someone might be drawn to both *Current Reflection* and the idea of inertness for essentially the same reason, namely, that subjects should respect their own opinions in the way captured by *Current Reflection*, which entails that learning about one's own opinion should not occasion a change in that opinion.⁷ Indeed, perhaps the idea behind inertness all along was merely that *learning* propositions about one's doxastic states cannot have

⁵ Not all of the problematic propositions are strictly about my own credences: a similar problem arises if we let p be the proposition that *some* subject assigns a credence of 0.8 to some proposition (assuming it is not already rational for me to be certain of this proposition).

⁶ Van Fraassen (1984: 248) himself, for instance, remarks that the synchronic version of Reflection should be uncontroversial.

⁷ See Christensen's (2007) criticism of Dutch Book arguments in favour of Current Reflection.

the relevant sorts of first- or higher-order import, where it is assumed that only true propositions can be learnt.

In light of this, friends of inertness might want to restrict the above theses further: they should only be taken to apply for those values r that one's credence in p in fact takes (assuming, again, that updating happens by conditionalisation). If *First-order inertness* is restricted in this way, then satisfying it will never force violations of *Current Reflection*. In fact, *Current Reflection* entails *First-order inertness* thus restricted. Then, any counterexample to the latter will also be a counterexample to the former. Indeed, my arguments below, targeted at *First-order inertness*, will also be arguments against *Current Reflection*.

I will now argue that the inertness thesis – even if restricted in the ways discussed above – are false.

III Against the inertness of evidence about one's own doxastic states

(i) Chandra's prediction

Though some version of the inertness theses that might sound plausible at first hearing, there are situations in which it is far from obvious that evidence about one's own doxastic states is inert. Consider, first, situations in which subjects have reason to think that their current credences are either systematically too high or too low:

Chandra's prediction 1

Chandra has spent his life predicting the outcomes of political elections. Based on a vast body of evidence E_{ORIGINAL} , he has formed a credence in the proposition that Obama will win the presidential race. An angel whom Chandra has every reason to trust tells him that he has a strong tendency to slightly under-estimate the prospects of candidates who are running for a second term in office – a tendency that, unfortunately, persists even upon being told about it by an angel. Chandra then learns that he is 80% confident that Obama will win.⁸

Consider the time (t_1) at which Chandra has received the testimony of the angel, but has not yet learnt about his own credence. We can assume that despite the trustworthiness of the angel, her testimony is misleading, and Chandra's credences are in fact perfectly rational. Still, since Chandra has evidence that his credence – whatever it is – is too low, learning of his own 0.8 confidence that Obama will win should, it seems, give him some evidence that the evidence he has at t_1 supports Obama winning to a degree above 0.8. Nor does it seem implausible that Chandra's credences should violate *First-order inertness*: upon learning that he is 0.8 confident

⁸ There are similar cases in the recent literature in which subjects seem to violate *First-order inertness*. Take, for instance, Christensen's (2007) example (which he gives as an argument against *Current Reflection*) of a subject who thinks she is overly optimistic about the weather. Also, Egan & Elga....

that Obama will win, shouldn't he raise his confidence, given that he has excellent reason to think that it is too low?⁹

Someone might object that disapproving of one's own current credences as too low, in the way that Chandra does, is incompatible with being perfectly ideal. But it would be question-begging to simply appeal to a principle like *Current Reflection* or the inertness theses formulated above in support of this idea. Moreover, at least to show that *Higher-order inertness* is incorrect, we can, instead, consider a case in which a subject acquires evidence that his credences are, at least within a limited domain, perfectly ideal:

Chandra's prediction 2

Chandra has spent his life predicting the outcomes of political elections. Based on a vast body of evidence E_{ORIGINAL} , he has formed a credence in the proposition that Obama will win the presidential race. An angel whom Chandra has every reason to trust tells Chandra that when it comes to propositions about the outcomes of the election, Chandra's credences are (and will continue to be) ideal. At this point Chandra is neither sure what his own credence is, nor what the ideal credence is, in Obama winning. He then learns that he is 80% confident that Obama will win.

Assume that the angel is in fact right: Chandra's credence in Obama winning (both initially, and after the testimony of the angel) is ideal. Assuming that Chandra has very strong reason to trust the angel, conditional on his credence in Obama winning being some value r , he is virtually certain that r is ideal. However, before learning about his own confidence he is not at all certain that a confidence of 0.8 in Obama winning is ideal. In this case it seems that learning about his own credence could *at least* give Chandra evidence about what the ideal, rational credence is that Obama will win.¹⁰

Finally, consider a more mundane case in which Chandra merely regards his doxastic states as reliably, though not infallibly, reflecting his evidence. Perhaps he knows how reliable he is through the testimony of the angel, perhaps he considers his own track record, or perhaps his views about his own reliability have some other rational basis. If learning about his own doxastic states can at least have higher-order evidential import when he knows, or is certain, that his credences are correct, it would seem bizarre if they couldn't do so when Chandra merely thinks that it is

⁹ Letting o be the proposition that Obama will win, and Cre_t be Chandra's credence function at time t just before he learns of his own 0.8 confidence in o , we can, for instance, assume that $\text{Cre}_t(o \mid \text{cre}_t(o) = 0.8) = 0.875$. There are numerous credence distributions Chandra might have concerning what his own credence in o is that would make him probabilistically coherent. Here is one: Chandra is 0.9 confident that his credence in o is 0.8 (the credence he actually has), and 0.1 confident that his credence is 0.114. Moreover, because he takes himself to systematically under-estimate, $\text{Cre}_t(o \mid \text{cre}_t(o) = 0.114) = 0.125$. Then, $\text{Cre}_t(o) = \text{Cre}_t(o \mid \text{cre}_t(o) = 0.8) \times \text{Cre}_t(\text{cre}_t(o) = 0.8) + \text{Cre}_t(o \mid \text{cre}_t(o) = 0.114) \times \text{Cre}_t(\text{cre}_t(o) = 0.114) = 0.875 \times 0.9 + 0.125 \times 0.1 = 0.8$.

¹⁰ Below I consider a view on which it can give Chandra evidence about the ideal credence, but not about who will win the nomination.

likely that his credences are correct.¹¹ The idea that evidence about one's own doxastic states is inert is already beginning to look a lot less intuitive.

(ii) Some puzzling consequences

One of the nefarious activities of Madame Babineaux is exploiting the linguistic ineptitude of vegetarian tourists by making them order bone marrow soup in her Parisian bistro. My friends and I did our best to place our order, and to explain what a vegan is, and the Madame just repeated, amid the very loud chatter, what dish she has us all down for. We also have various other clues as to whom she is about to play her trick on. Let E_{ORIGINAL} be our common body of evidence about what dish it is that she is about to send our way, and b the proposition that we are about to receive some of the infamous soup. At a time t_0 I have done my best to determine whether or not b is true, but unfortunately, I have no better access to my own opinion than I do to the opinions that others might hold about the issue.¹² In fact, the evidence makes b likely (say to degree 0.9).

I am then, at a slightly later time t_1 , given the following information: one of my peers is confident, based on the original evidence, that b is true, whereas another is confident that b is false. Call these Pro-peer and Con-peer. My total evidence now consists of something like the following:

E₁:

- E_{ORIGINAL}
- Pro-peer is confident that b , whereas Con-peer is confident that $\neg b$.

Now, I take it that those who subscribe to the idea that learning about the opinions of peers can provide evidence both about one's evidence and about the world want to say that in the kind of situation described I (at least may) acquire some evidence that my original evidence doesn't support b , and that on the total evidence I now have it is reasonable to be at least a bit less confident in b than based on E_{ORIGINAL} alone.

But finally, at a yet later time t_2 I learn that I am Con-peer. My total evidence is now:

E₂:

- E_{ORIGINAL}
- Pro-peer is confident that b , whereas I am confident that $\neg b$.

¹¹ We might also consider an intermediate case: if the angel tells Chandra that he has a perfectly ideal credence, but Chandra thinks there is a slight (10%) chance that the angel is lying, then upon learning, for instance, that his credence that Obama will win is 0.95, shouldn't Chandra be roughly 90% confident that this credence is ideal? How does this differ from a case in which Chandra merely regards his own credences as 90% reliable, and then learns of his 0.95 credence?

¹² I take this to be compatible with the thought that I may have different kinds of evidence about my own opinions and the opinions of others; it's just that I don't have *better* evidence about my own opinion.

If first-person evidence is inert, then as far as the relevant propositions go (propositions about how likely b is on my evidence, and b itself), learning that I am Con-peer should completely wipe out the effect of learning that Con-peer is confident that b is false. It is as though my evidence now consisted of just the following:

E₃:

- E_{ORIGINAL}
- Pro-peer is confident that b .

Now, I take it that those who think that evidence about the opinions of others is not inert will agree that **E₁** and **E₃** make reasonable different degrees of confidence in the relevant propositions: **E₁** contains misleading evidence about my original evidence, whereas **E₃** does not.¹³ It follows that *merely* learning that I am Con-peer should make me boost my confidence in b . Now, this in itself may seem bizarre: why should merely learning that I am Con-peer suddenly render Con-peer's opinion inert? What justifies refusing to give my own judgments any weight at all?

Here is a case in which the inertness theses have an entailment that seems even more bizarre. Assume that I learn that Con-peer was just duplicated, and I am in fact her molecule-for-molecule duplicate. As a result, I know that now, at t_0 , I hold exactly the same opinions about whether b as Con-peer does. If evidence about my own doxastic states is inert, then learning that I am confident that $\neg b$ should have no evidential relevance for what the original evidence supported, or for whether b . But if a proposition p is inert, and I am certain that p is true just in case a proposition q is, then q ought also to be inert. Because I am certain that I am confident that $\neg b$ if and only if Con-peer is confident that $\neg b$, it follows that learning that Con-peer is confident that $\neg b$ must similarly be inert. But this is very bizarre: why should merely knowing that I am a duplicate of Con-peer completely screen out the evidential relevance of learning about her opinions?¹⁴

Returning to the original case in which I learn that I am Con-peer, it is also worth noting that the kind of erasing of the force of Con-peer's opinion entailed by the inertness theses would go against popular verdicts about cases of peer disagreement. According to many who hold so-called conciliatory views, I should give at least some weight to my own opinion – indeed, Elga's (1997) Equal Weight View, for instance, urges giving my own opinion and that of my peer equal weights, and what White (2009) dubs the “thermometer model”, I should treat my own credences as a guide to what is true in exactly the same way that I treat your credences. But consider the fact that the evidential situation I am in when I have evidence **E₂** is a somewhat standard peer disagreement case. According to a popular

¹³ This is so even if I initially took into account the (say slight) possibility that I am either Con-peer or Pro-peer. In any case, in so far as I have no reason to think that I am one rather than the other, it seems that whatever discounting of one opinion I do, I should equally discount the other.

¹⁴ I am indebted to Billy Dunaway for bringing up the issue of duplication, which helped me think of this example.

view, in such cases I ought to assign at least some weight to both opinions, ending up with a confidence in b that is somewhere between that of my own and of my peer.¹⁵ By contrast, if being confident in b is rational on the original evidence, and I then learn that my peer is confident in b , surely this won't make it rational for assign to b a slightly lower credence. Hence, given standard "conciliatorist" views of peer disagreement, evidence \mathbf{E}_3 and evidence \mathbf{E}_2 make reasonable different attitudes to b (and to propositions about which credence in b is rational). Given that \mathbf{E}_2 differs from \mathbf{E}_3 only in that I know what my opinion is, it follows that that opinion cannot be evidentially inert. Hence, the thesis of the inertness of first-person evidence appears to be incompatible with popular views about peer disagreement.

Someone might respond that those writing on peer disagreement simply assume some form of luminosity. Then – the thought goes – it is impossible for a subject who has formed a credence in b to have *just* body of evidence \mathbf{E}_3 . But even if assumptions are made entailing that a subject could never have evidence \mathbf{E}_3 , we can still ask what that evidence supports. And the problem is that if my own opinion (of which I have trivial knowledge) was genuinely inert, then \mathbf{E}_2 and \mathbf{E}_3 should lend exactly the same degree of support to b . But as we have seen, on somewhat standard conciliatorist commitments, they do not. Hence, it is difficult to see how luminosity assumptions could let the conciliatorist who thinks that first-person evidence is inert completely off the hook.¹⁶

I will now argue further that at least if I am a conditionaliser, regarding myself as above 50% likely to assign the rational credence to b , whatever that credence is, entails at least the falsity of *Higher-order inertness*. This will conclude my initial attack of the inertness theses.

(ii) A Bayesian argument

Assume that having evaluated a body of evidence $\mathbf{E}_{\text{ORIGINAL}}$, I am not certain what the import of my evidence is regarding a proposition p (for instance, the proposition that we are about to receive some bone marrow soup). I distribute my credences in its import among a partition of hypotheses about how likely my evidence (objectively) makes p at time t , or what the rational credence in p is given that evidence. Let $\{p_o(p) = r_1, \dots, p_o(p) = r_n\}$ form such a partition, and let $\text{Cre}(-)$ be my

¹⁵ See [Elga, Kelly](#), White (2009) etc... Christensen (2011) is an exception, for he defends the idea that evidence about my own states is inert, and recognises that (i) this entails that \mathbf{E}_2 and \mathbf{E}_3 support the relevant proposition (in this case b) to the same degree, and that (ii) \mathbf{E}_3 doesn't, it seems, support b to a lower degree than $\mathbf{E}_{\text{ORIGINAL}}$ alone. Hence, he endorses the conclusion that the rational attitude to b based on \mathbf{E}_2 is no lower than that based on just $\mathbf{E}_{\text{ORIGINAL}}$. However, this is not a conclusion that most conciliatorists endorse.

¹⁶ In the example given it was assumed that my original opinion was in fact irrational. As Christensen (2011: 4) points out, conciliatorists are not committed to saying that if a subject takes evidence about peer disagreement into account as they recommend, she automatically ends up with a rational opinion. However, even if we take conciliatorism to be a view about how a particular kind of evidence (i.e. evidence about peer disagreement) ought to be taken into account, the point made shows conciliatorism to be incompatible with inertness of first-person evidence.

credence function at time t .¹⁷ Assume further that not only do I regard myself as reliable in the sense of being likely to evaluate my evidence correctly, but I consider myself likely to form the correct opinion *no matter what the correct opinion is*. In other words, for any hypotheses $p_o(p) = r_i \in \{p_o(p) = r_1, \dots, p_o(p) = r_n\}$ about the degree to which my evidence supports p , conditional on $p_o(p) = r_i$ obtaining, I consider myself likely at least to some degree above 0.5 to assign the credence in accordance with that hypothesis – i.e. the correct credence. Then, with ‘V’ for disjunction, the assumptions made are the following:

1. $\text{Cre}(p_o(p) = r_i | p_o(p) = r_i) > 0.5$
2. $\text{Cre}(\bigvee_{1 \leq i \leq n} p_o(p) = r_i) = 1$, and for all i , $0 < \text{Cre}(p_o(p) = r_i)$

These assumptions entail that for arbitrary $i \in \{1, \dots, n\}$

3. $\text{Cre}(p_o(p) = r_i | \text{cre}(p) = r_i) > \text{Cre}(p_o(p) = r_i)$.¹⁸

Hence, as long as my credences satisfy 1. and 2., they will fail to comply with *Higher-order inertness*. As long as I conditionalise, for any i , learning that my credence in p is r_i should make me *raise* my credence that my original evidence supported p to degree r_i .

To say the least, it certainly seems possible for the above assumptions to hold for a rational subject. Moreover, 1. doesn’t look like an implausible condition on what it is for a subject to regard herself as a reliable evaluator of the evidence. Consider the following simple case. I am not sure to what extent my evidence supports p , but I do know that either it makes p likely, or it makes p unlikely (perhaps this much has been revealed to me by an epistemology oracle). I am not sure what I myself think about the matter, but I treat my credences as tracking my evidence with over 50% success: conditional on my evidence making p likely, I regard myself as over 50% likely to have a high degree of confidence in p , and conditional on my evidence making p unlikely, I regard myself as over 50% likely to have a low degree of confidence in p . It follows that merely learning that I am confident (unconfident) in p should make me boost my confidence that my evidence supports p to a high (low) degree.

Regarding myself as a reliable evaluator of my evidence seems to force violations of *Higher-order inertness*, but what about *First-order inertness*? Let two propositions be *strongly dependent* if any learning that changes the probability of

¹⁷ I assign a non-zero credence to each hypothesis in the partition. One might worry that there are uncountably many such hypothesis, but to avoid such problems, we need not think about each r_i as a point value; I leave open the possibility that these are intervals. What is important is just that r_1, \dots, r_n are disjoint. If we think of the members of the partition as intervals, then talk about assigning a credence of r_i to a proposition should be understood in terms of having a credence that falls within the interval r_i . **Moreover, we can even avoid the assumption that there is a uniquely rational credence: the members of the partition might be conjunctions of rational credences. In this case, the argument below would have to be slightly modified.**

¹⁸ The proof is in the Appendix.

one also changes the probability of the other. Then, assuming that updating happens by conditionalization, two propositions p and q are strongly dependent just in case $0 < \text{Cre}(p) < 1$, and for any e

$$\text{Cre}(p|e) \neq \text{Cre}(p) \text{ iff } \text{Cre}(q|e) \neq \text{Cre}(q).^{19}$$

Below I argue that sometimes propositions about what the ideal or rational credence in a proposition p is and p itself are, indeed, strongly dependent. In so far as the above assumptions hold in some such cases, it follows that sometimes learning of your own credence in p may make it rationally mandatory to change that credence.

Having presented a tentative case for the falsity of the inertness theses, I will now discuss a way of trying to re-formulate an asymmetry between the evidential import of third- and first-person doxastic states that acknowledges the arguments given above. This will also allow me to complete my argument against *First-order inertness*.

IV A refined version of first-order inertness

Perhaps the arguments given above still leave room for an interesting asymmetry between the epistemic import of evidence about one's own doxastic states and evidence about the doxastic states of others. In particular, there is something particularly implausible about the idea that learning about my own perfectly rational credence in a proposition p could make it rational for me to change my credence in p . Consider, in particular, the following kind of case:

First-person boost

Based on evaluating a body of evidence E_{ORIGINAL} , I become fairly confident that we are victims of Madame Babineaux's nefarious schemes (and hence, of proposition b). I know that I am reliable at evaluating such evidence. Upon learning of my own confidence in b , I further boost my confidence, becoming very confident that the Madame is about to play her trick on us.

Assume that I have no reason to think (as Chandra did in one of the cases described above) that I tend to under-estimate the force of my evidence. If this is so, isn't it deeply problematic that I could boost my confidence in a proposition just by discovering that I am confident that it is true?

Perhaps this, then, is at the heart of the putative asymmetry: despite the fact that I am perfectly rational, learning what you think about p can still make it rational for me to change my opinion about the matter, but if I am perfectly rational, and have no reason to think that my credences are either systematically too high or too low, then learning what I think about p cannot make it rational for me to change my

¹⁹ Hence, a simple example of strongly dependent propositions are p and $\neg p$, when one is certain neither that p is true nor that it is false.

opinion. The idea is that though *Higher-order inertness* fails, *First-order inertness* holds given the (further) restrictions that I am rational, and have no reason to think that I am prone to err in one direction rather than another. It is not obvious that anything has been said so far that would constitute an argument against this new, refined version of the inertness thesis.²⁰

What might motivate the view that though learning about my own rational credences can have higher-order import, it cannot have first-order import? I can only think of one train of thought leading to such a view. Assume that I have a perfectly rational credence of 0.9 in a proposition p . On the refined inertness view, if I regard myself as reliable, then learning that I am 0.9 confident in p may be evidence that 0.9 is the ideal, rational credence. I should take some credence away from hypotheses stating that the rational credence is something other than 0.9, and move all that credence to the 0.9 hypothesis. But assume, first, that (i) I take credence away from these other hypotheses in such a way that their proportions remain the same. (And shouldn't I do precisely this if I have no reason to think that I tend to err in one direction rather than the other, being prone to either over- or under-estimate the import of my evidence?) And assume, second, that (ii) my initial credence of 0.9 in p was my expectation of the correct, ideal credence. Then, even if I boost my confidence that 0.9 is the ideal credence, my *expectation* of the ideal credence will remain at 0.9. But (iii) if my expectation of the ideal credence doesn't change, then whatever evidence I acquired cannot have relevance for whether p . The same kind of argument could be applied to cases involving evidence about the doxastic states of others: if I start out with an ideal credence of 0.9 in a proposition p , then learning that your credence is 0.9 shouldn't effect my credence in p , as long as I have no reason to think that you tend to either over- or under-estimate the import of the evidence.

One of the core assumptions of the above argument is that my credence in a proposition should always equal my expectation of the ideal credence:

Rational Expectation

$$\text{Cre}_1(p) = \text{Exp}_1[p_o(p)]$$

This principle is entailed by a principle that I will (following David Christensen²¹) call *Rational Reflection*:

Rational Reflection

$$\text{Cre}_t(p \mid p_o(p) = r) = r$$

Hence, we have an argument for a refined thesis of first-order inertness employing *Rational Expectation* (or *Rational Reflection*) as a core assumption. The problem with the argument is that such principles place what strike me as implausible

²⁰ For instance, in pointing to the bizarre consequences of learning certain identity-facts, like the fact that I am Con-peer, I was assuming that it was the opinion of Pro-peer that was rational to start out with.

²¹ Christensen (2010).

restrictions on the kinds of evidential situations that it is possible for subjects to be in.

As an example, consider the case of *Clock beliefs*.²² You are looking at the minute hand of an unmarked clock from some distance away. The hand moves in discreet one-minute jumps. Given your perceptual abilities and your distance from the clock, you are not an infallible judge as to the exact position of the hand. For this reason, even though the hand (let us assume) in fact points to 20 past the hour, given the nature of your perceptual evidence, you shouldn't be completely certain that this is the case. Assume, for instance, that you should assign some credence to the hand pointing anywhere between 18 to 22 past the hour. Assume also that (given your distance from the clock) something similar is true in other cases: you always have a 2-minute "margin for error" in either direction (and we can assume that you know this). So, for instance, conditional on the hand pointing to 18 past the hour, it is rational for you to assign some credence to its pointing anywhere between 16 and 20 past the hour.

Recall that the hand in fact points to 20 past the hour. For simplicity I will assume that you assign equal credence to all five relevant hypotheses²³. Let h be the following proposition:

h : The minute hand points to either 19, 20, or 21 past the hour.

In the model described your credence in h ought to be 0.6. However, it is not certain that this is the rational credence in h : for instance, conditional on the hand pointing to 18 past the hour, the rational credence in h is only 0.4. It follows that the credences described violate *Rational Expectation* and *Rational Reflection*. For consider the fact that if the hand points to either 18 or 22, the rational credence in h is 0.4, and if any of the other cases obtain, the rational credence is 0.6. Hence, the rational expectation of the ideal credence is:

$$0.6 * 0.6 + 0.4 * 0.4 = 0.52 < 0.6$$

In so far as this kind of evidential situation is possible, *Rational Expectation* fails, and so does the above defence of the refined inertness thesis.

In *Clock beliefs*, which credence it is rational to assign to propositions about the position of the minute hand (propositions such as h) depends on what the position of the minute hand in fact is. But not only this, there are pairs of first- and higher-order propositions that are *strongly dependent* in the sense introduced above: both are dependent on some propositions, and they are dependent on exactly

²² Adapted from David Christensen's (2010) example, which is adapted from Williamson ().

²³ Perhaps one ought to assign a higher credence to the hand pointing to 20 past the hour than to other hypothesis, and similarly, a higher credence to its pointing to 19 (or 21) than to its pointing to 18 (or 22), but problems for *Rational Reflection* are created whether you distribute credences equally among the hypotheses or not.

the same propositions.²⁴ For instance, h and $p_0(h) = 0.6$ are strongly dependent, for it is rational to assign to h a credence of 0.6 just in case h is true. The reason why this is interesting in the present context is that it shows that there are cases in which any failure of *Higher-order inertness* will *inevitably* lead to a failure of *First-order inertness*: in such cases changing one's confidence in higher-order propositions about which credence in some proposition p is rational will force one to change one's confidence in p .²⁵ So, for instance, if I regard my own credences as reliable in the sense specified in connection with the Bayesian argument given above, and I boost my confidence that 0.6 is the rational credence as a result of learning that my own credence in h is 0.6, I should also boost my confidence in h . In light of this, it is far from clear whether there is any interesting, viable refinement of *First-order inertness*.

It is worth briefly noting that one solution to the kinds of problems that cases like *Clock beliefs* create for *Rational Reflection* is to modify the principle. In particular, one might suggest modifying it in a way that follows Ned Hall's suggested modification of Lewis' Principal Principle.²⁶ The thought is that you ought to adopt the opinions of experts who have all the evidence you have. However, learning that a certain credence was rational at a time t does not entail that it is still rational, as now you have more evidence than you did at t – after all, you now know this fact about what was rational at time t . The idea is:

New Rational Reflection

$$\text{Cre}_t(p \mid P' \text{ is ideal}) = P'(p \mid P' \text{ is ideal})^{27}$$

Here Cre_t is the credence function of a rational subject s , P' is a credence function, and “ideal” means “the ideal, rational credence function for s to have given her evidence at t ”. I cannot here give this new principle the discussion it merits²⁸, but only wish to note that it cannot be used to support the refined thesis of first-order asymmetry in any obvious way, since it does not entail that when a subject's expectation of the rational credence in a proposition p doesn't change, neither should her credence in p .

Despite everything that has been said, one might still have various worries about a view on which evidence about one's own doxastic states has the kind of import I have argued for. I will now address some of these.

²⁴ Strong dependence entails dependence, but not vice versa, since (probabilistic) dependence is not a transitive relation: if p and q are dependent, and p is dependent on e , it does not follow that q is dependent on e .

²⁵ This is not to say that such strong dependence is necessary for failures of *Higher-order inertness* to lead to failures of *First-order inertness*; only that it is sufficient.

²⁶ See Hall () for a modification of Lewis' () principle, and Elga () for an analogous modification of *Rational Reflection*.

²⁷ See Elga (*manuscript*).

²⁸ I criticise the principle in work in progress.

V Some remaining objections

(i) The objections from double counting and informational proxies

Recall the case *First-person boost* in which I increase my confidence in a proposition upon learning that I am fairly confident that it is true. I have heard several people react to such cases by saying that they involve some illegitimate double counting or recycling of one's evidence.²⁹ What goes on, according to this objection, is not too different from the kind of flaw I commit in the following cases:

The thermometer

My belief about the temperature is based on looking at a thermometer that clearly reads -15°C . I then look at the thermometer again, and boost my confidence that the temperature outside is -15°C . I am aware of having looked at the very same thermometer twice, and the interval between the two observations is so small that I consider it certain that the thermometer will give the same reading on both occasions.

Laura and Dave

My belief about the temperature outside is based on Laura testifying that it is -15°C . I know that Dave repeats whatever Laura says about the weather. Nevertheless, upon hearing Dave say that it is -15°C , I boost my confidence that the temperature outside is -15°C .

Strictly speaking, at least the second case doesn't involve counting the very same evidence twice. Rather, I count as relevant a piece of evidence that I know to be thoroughly dependent on evidence that I have already taken into account. This is what gives rise to the charge that I am double counting: I am more or less recycling or counting twice whatever evidence is provided by my first observation of the thermometer, or by Laura's testimony. Matters would be different if I observed two different thermometers, or heard two testimonies that were suitably independent.

What is the analogy between the above cases and *First-person boost*? Consider first a simplified case in which all my evidence is perceptual, and I know that I believe a proposition p just in case it perceptually seems to me as if p . Then, my belief that p more or less just encodes how things perceptually seem to me: the perceptual seeming is like Laura's testimony, and the belief like Dave's. Here is a more general thought. In cases like *First-person boost* my original credence is based on something like how things evidentially seem to me (think of a perceptual seeming as a special case of an evidential seeming). As such, my credence merely encodes or represents that seeming. If I give evidential weight to my credence, I am in effect counting the seeming – and hence, my original evidence – twice. To make what is going on more vivid, imagine that there is a little homunculus in my mind. Whenever I am in a new evidential situation, the homunculus makes a

²⁹ Kelly (2005) mentions the idea that even counting the doxastic responses of *others* as more evidence involves an illegitimate sort of double counting.

recommendation, and my credences basically just encode that recommendation. If I then adjust my opinion as a result of learning of my own credence, it is as if I was counting the recommendations made by the homunculus twice, and in so doing treating the opinion of a single informant as the opinions of two independent informants.

The above assumes that evidential seeming (or the recommendations made by the homunculus) must count as part of my evidence in *First-person boost*. But even setting aside the question of whether the resulting view of evidence is at all plausible, the assumed analogies are dubious. There is a problem both with the assumption that my credences straightforwardly encode the evidential seemings (the recommendations of the homunculus), and with the assumption that the import of these seemings is always clear to me. Let me begin with the first point.

Just as it was assumed in *Laura and Dave* that Dave repeats whatever Laura says about the weather, it was assumed that my credences simply reflect my evidential seemings, and that I am practically certain of this. But in general, the relationship between items of evidence and credences cannot be so tight: even when E is part of my evidence, I may not know (or be certain) that E is true just in case my credences are a certain way.³⁰ If I did, then assuming that I know or am certain of items of my evidence, I should already know what my credences are: if I am certain of E , and I am certain that my credences are such-and-such just in case E , then I ought to be certain that my credences are such-and-such.³¹ Of course, this would conflict with the assumption made in *First-person boost* that I learn about my own credences. And there is also a further point to be made: if the homunculus model held, then it would not be clear how evidence about the doxastic states of *others* could have the kind of relevance we are after – or indeed, how cases of peer disagreement could arise in the first place. If I know that my friend and I have the same evidence, which amounts to both of our homunculi saying the same thing, and if there is no epistemic gap between credences and the testimony of the homunculus, then it wouldn't even be possible for us to disagree.

Now, one might object that the disanalogy pointed to is neither here nor there, for even if Dave only repeated what Laura said 80% of the time, tossing a coin about what to say the remaining 20% of the time, Dave's testimony could not give one new evidence about the temperature outside, and the charge of double-counting could still be mounted. Similarly, we need not assume that credences perfectly track evidential seemings. However, there is a further problem with the above analogies, which brings me to the second point. It was assumed that it is always clear what the import of the evidential seemings is, or what the homunculus testifies to. In the case

³⁰ If evidence consisted, instead, of propositions about the external world, such a claim would be very bizarre indeed. But for the sake of argument I am assuming that evidence consists of something like propositions about how things seem to one to be.

³¹ This follows straightforwardly from the probability axioms: if $\text{Cre}_t(E) = 1$, and $\text{Cre}_t(E \text{ iff } \text{cre}_t \text{ is such-and-such}) = 1$, then $\text{Cre}_t(\text{cre}_t \text{ is such-and-such}) = 1$. If instead of certainty we prefer to talk of knowledge, then assuming closure we should still get something close to the conclusion that I am at least in a position to know what my own credences are. This is still problematic, in so far as my lack of knowledge of my own credences in *First-person boost* is not merely a matter of not having performed the right logical inferences.

described above, Dave's testimony doesn't yield any evidence about the import of Laura's testimony, or about exactly what it was she said. But even if Dave repeats what Laura says, this need not always be the case: if Laura shouts into the wind and is barely audible, Dave's testimony might well have evidential relevance. In the present context, assuming that learning about one's doxastic states cannot yield evidence about the import of one's original evidence begs the question, since that is precisely what is at issue in the debate over higher-order inertness. If there is uncertainty about that the testimony of the homunculus, and one's credences are very reliable at tracking it, then some further argument must be given for the conclusion that learning about those credences couldn't give evidence about it. The charge of double counting only applies if precisely what is at dispute is already assumed.

What has been said also defuses a related objection:

"Belief-states are a kind of informational proxy for one's evidence, just as transcripts of documents are proxies of originals, or pictures of scenes are proxies of those scenes. The evidential force of such proxies of imperfect representations is screened out by the things they are representations of".³²

Sometimes a transcript of a document misrepresents the original. And even when it doesn't, any information gained by reading it can also be gained by reading the original. Either way, in the presence of the original document the transcript is evidentially inert – or so goes the objection.

However, it is not true that the evidential force of a proxy is always swamped by the thing it is a proxy for. For instance, assume that an original document is written by hand, and that it is difficult to make out the writing. The transcript, by contrast, is typed. You know that the person who typed it was reliable, though far from infallible, at deciphering hand-written documents. Hence, their transcript is only an incomplete representation of the original. Nevertheless, the transcript can provide evidence that is not screened off by the presence of the original document: your access to what the original document says is limited, and the transcript is a source of information about what the original says.

The objection from informational proxies has an even more obvious problem, namely, that it generalizes too far: if a doxastic state is a proxy that loses any evidential force it might have in the presence of the evidence it is a proxy for, then information about other subjects' beliefs would be evidentially impotent. But we were looking for a problem that is specific to one's own doxastic states.³³

(ii) The objection from single-sidedness

Learning about the opinions of others can provide a genuine check on one's own opinions: the rationality or correctness of one's opinions is confirmed when others

³² Kelly () discusses an argument along these lines...

³³ Cf. Kelly

agree, and disconfirmed when they disagree. One worry is that evidence about one's own doxastic states doesn't have this dual aspect, and for that reason, cannot function as genuine evidence³⁴:

“If I am confident in p , then learning of my own confidence is evidence that my evidence supports p , as well as evidence that p is true. But had I been confident in $\sim p$, learning about my own confidence would have provided me with evidence that my evidence supports $\sim p$, as well as evidence that p is false. So learning about my own doxastic states could only ever give me evidence confirming the correctness of those states, but never evidence disconfirming their correctness. But surely it is a *criterion* of proper evidence that it can confirm or disconfirm the correctness of an opinion.”

There appear to be several issues that this objection is running together.

It has, indeed, been proposed that if the outcome of a test or procedure is to be evidentially relevant for some hypothesis, then it must be capable of both confirming and disconfirming that hypothesis.³⁵ But whether or not this is right, it cannot be used to discredit cases like *First-person boost*, for it is not as if there is any one hypothesis, whether about the degree to which my evidence supports a proposition p , or about p itself, that is confirmed no matter what I learn my doxastic states to be. Moreover, nothing I have said entails that I should become more confident of the proposition that my opinion (whatever it is) is correct merely as a result of learning what that opinion is: it is one thing to become more confident that a credence of 0.9 in p , for instance, is correct, and quite another to become more confident that my credence in p (whatever it is) is correct.

Nevertheless, part of the above objection was that learning about my own credences cannot provide a check on those credences, for it cannot disconfirm the proposition that my opinion is correct, whereas learning about the credences of others can do so. Now, this is not quite right: recall the Chandra cases above. But more importantly, even when my situation is set up so that no such disconfirming of the rationality of my own opinion can occur, this hardly supports the thesis that evidence about one's own opinions is inert. Just as learning about my own opinions doesn't typically disconfirm the rationality or correctness of those opinions, learning about the opinions of others doesn't typically disconfirm the correctness of their opinions. It was no part of my claim that evidence about one's own opinions and evidence about the opinions of others must confirm and disconfirm exactly the same propositions. Rather, the claim was that both kinds of evidence can confirm

³⁴ For instance, Christensen (2011: 7) says something along these lines: “we may take the first-person psychological evidence to be incapable of providing the sort of check on one's reasoning that third-person evidence provides. In this sense, it is relatively inert”.

³⁵ This has been one reason to discredit “bootstrapping” – reasoning to the reliability of one's faculties from premises that one knows or justifiably believes based on employing those very faculties – on the grounds that bootstrapping can only ever confirm the reliability of my faculties. See Pryor, White, etc...

(and disconfirm) propositions about what one's evidence supports, and propositions about first-order matters.

(iii) The objection from the Transparency Thesis

Assume that (perhaps contrary to fact) there is a strong correlation between rain and the BBC weather forecast predicting rain. Then, based on seeing that it is raining, it may be perfectly rational to increase my confidence that last night's BBC forecast predicted rain. But now assume that I come to know that it is raining (partly) based on the forecast. In this case, I cannot use this knowledge to increase my confidence that the BBC forecasted rain – that would involve a problematic sort of epistemic circularity. Neither, it might seem, should I increase my confidence that, whatever was said on the BBC forecast, it strongly supported the claim that it would rain. In the first case I would be relying on some evidence E to come to know a proposition p , and then employing p to boost my confidence in E . In the second, I would be relying on E to come to know p , and then employing p to boost my confidence that E supports p .

Similarly, if I base my belief that I believe p on p itself, then I cannot use the fact that I believe p as further evidence for p . Or, if I base my belief that I believe p on my first-order evidence for p , then perhaps I cannot use the fact that I believe p as further evidence about that evidence. But if some form of the Transparency Thesis is true, then the worry arises that cases like *First-person boost* involve some such epistemic circularity. According to defenders of the thesis, when I wonder what my mental state regarding a proposition p is, I attend to precisely the same evidence or considerations as those I would attend to when I wonder whether p :

“The typical way in which subjects come to know what their doxastic states are is not by some mechanism of looking inwards, but by looking outwards: the evidence on which I come to know that I believe that p is the very same evidence as that on which I come to know p . If I base my belief that I believe p on a body of evidence E , I cannot then employ my knowledge of my belief as a basis for further beliefs about E or p .”³⁶

Here is a way of defending this idea within a credal context. Assume that there is a reliable connection between my evidence supporting a proposition to some degree and me assigning to that proposition a credence reflecting that degree of support. The thought is that then, if my evidence E in fact supports p to degree 0.9, believing,

³⁶ The idea of transparency is normally attributed to Gareth Evans (). I am considering the defence given by Fernandez (2003). The reason I focus on this defence is first, that it can be modified to explain first-person access to more fine-grained doxastic states. For instance, Dretske's suggestion (that Byrne's () view, for instance, is very close to) that I infer that I believe p from my awareness of p doesn't have any obvious application to my knowledge of my degrees of confidence or credences. Second, the suggestion made by Fernandez doesn't fall prey to obvious problems encountered by some other ways of cashing out transparency. A common problem is saying how, of I want to know what my state is at a time t , a transparent mechanism can give me knowledge of my state at t , rather than a later time at which I have a more comprehensive body of evidence.

based on E , that my credence in p is 0.9 is a reliable way of forming beliefs about my own credences. But why exactly is this? E , let us assume, is a set of propositions about the external world. Surely there need not be any connection whatsoever between these propositions being true and my credences. Nevertheless, because there is a reliable connection between *having* E as evidence and believing p to a certain degree, and I couldn't base my belief on E unless I had E as evidence, this method of forming beliefs is reliable. However, it is only available to me: you cannot justifiably form beliefs about my beliefs just based on E .³⁷

First, even if some version of the Transparency Thesis was true, that would not undermine my main point: there is no motivation for thinking that evidence about the doxastic states of others can have certain kinds of evidential import, but evidence about one's own states can never do so. Even those who defend Transparency don't think that it is the *only* way of finding out about one's states. I could instead resort to the kinds of methods by which I come to know about the mental states of others, by observing my own behaviour – or a neuroscientist could simply tell me what I believe. If in such cases the above objection would have no bite.

Second, I am not at all convinced that any form of the Transparency Thesis is true. The above account of first-person access strikes me as problematic for numerous reasons. Assume that my credences track degrees of evidential support fairly reliably, but only 80% of the time. It is far from intuitive that such correlations would be enough to yield *knowledge* of one's own mental states by the above kind of reasoning. Still, does privileged access really dwindle with the reliability of one's doxastic states? It is far from obvious that it does. Further, the kind of reasoning given above seems to me to generalise too far. Consider the fact that there is (trivially) a very strong correlation between having E as evidence and having E as evidence. Then, it believing that my evidence consists of all and only E based on my total evidence E would be a very reliable belief-forming process, since I couldn't, again, base a belief on E if I didn't have E as evidence. However, sometimes subjects are in no position to know just what their evidence is. But how can this be reconciled with the idea that a subject is always entitled to believe that her evidence consists of E based on E ? One might reply that though this entitlement exists, sometimes subjects don't have enough access to their own evidence to be in any position to base beliefs on that evidence. But then this should apply to beliefs about one's own doxastic states as well. Does first-person access really depend on contingencies about one's access to the evidence?

At any rate, even the Transparency Thesis would not establish that evidence about one's own opinions is inert. At most, it would establish that it is inert a lot of the time, and much more often evidence about the opinions of others. But even this would depend on a viable version of Transparency.

³⁷ See Fernandez (2003: 362-363).

(iv) The objection from bootstrapping (*warning: this is still a bit sketchy*)

One worry about the view I have been defending is that if I can boost my confidence in a proposition as a result of learning how confident I am in it, then this leads to the absurd consequence that I can bootstrap myself to a higher and higher confidence by repeating the same process.³⁸ Perhaps, through enough repetitions, I can work up to a confidence that is arbitrarily close to 1.

Assume that at a time t_1 I learn the proposition $\text{cre}_0(p) = r_0$, and that this proposition is positively evidentially relevant for p . The objection assumes that at t_1 the proposition $\text{cre}_1(p) = r_1$ must likewise be positively relevant for p . But whether or not evidence about one's doxastic states has first-order evidential import depends on specific features of one's epistemic situation, and the fact that my situation at one time has such features doesn't in any way guarantee that those features will continue to be present at subsequent times. It is no part of the view I am defending that evidence about one's own states *always* has either first- or higher-order import. Second, if $\text{cre}_0(p) = r_0$ and $\text{cre}_1(p) = r_1$ are genuinely two different pieces of information for me, it is not clear why their both being relevant for p would be so absurd after all. And finally, even if evidence about my own doxastic states continued to be positively evidentially relevant at subsequent times, it doesn't follow that I could eventually reach a confidence in p that is arbitrarily close to 1, for it may be that at each step I could increase my confidence less and less, and my credence started approaching some value below 1.

It is not difficult to construct cases in which $\text{cre}_0(p) = r_0$ is relevant for p at t_0 , $\text{cre}_1(-)$ results from conditionalising $\text{cre}_0(-)$ on information about one's credence in p at t_0 , but in which $\text{cre}_1(p) = r_1$ is not relevant for p at t_1 . A simple case of this sort is one in which upon learning $\text{cre}_0(p) = r_0$ at time t_1 , one also becomes certain of $\text{cre}_1(p) = r_1$. In such cases, the two propositions about one's credences at times t_0 and t_1 aren't genuinely two different pieces of information, for learning one is in effect learning both.³⁹ But even when one genuinely acquires two different pieces of information, there is nothing inevitable about such information being relevant for p on both occasions.

Return to the case of *Clock beliefs* described above. Assume that in such cases learning how confident one is in h , a proposition about the position of the minute hand at a time t_0 , can have evidential import for h . Recall the strong dependence between h and propositions about which credences in h are rational (at t_0). Because facts about what it *was* rational to believe at t_0 , given various positions of the hand at that time, don't change, the strong dependence continues to obtain at subsequent times, including t_1 . Then, we can answer the question about whether subsequently acquired evidence about one's own credences (in particular, evidence about one's credences at t_1 , acquired at time t_2) provides evidence about h by asking whether such evidence provides any new evidence about what it was rational to believe in the original clock scenario at t_0 . If it does not, then neither can it provide any new evidence about h .

³⁸ Thanks to Brian Weatherson for making me think about this objection.

³⁹ *Example...*

Assume first that I update my credences in the following sequential fashion: I take my old doxastic state as a starting point, and simply update it on the new evidence I have acquired, not going back to re-consider evidence I have already updated on. Moreover, I know that I update in such a sequential fashion.⁴⁰ Further, assume that at time t_2 I don't merely learn my credence in h , but I actually learn what my entire doxastic state or credence function is at a slightly earlier time t_1 . Bearing these assumptions in mind, consider the situation I am in just before I learn what my credences are at t_1 . Assuming evidential relevance to be symmetric, this first-person evidence could only be relevant for what it was rational to believe in the original clock scenario at t_0 if evidence about what it was rational to believe at t_0 was relevant for what my credences are at t_1 . But I know (we are assuming) just what my credences were at t_0 , and further, I know that my credences at t_1 are a result of updating those credences on the information about my credences at t_0 . Given these assumptions, it is not clear why information about whether my credences were in fact rational at t_0 would give me any information about what my credences are at t_1 . In fact, on what seems to me to be the most natural way of construing the case, it would not.⁴¹

Nevertheless, it is not at all implausible that subsequent evidence about my own doxastic state *could* provide more evidence about h . Consider first cases in which I go back to re-consider a body of evidence: knowing that the minute hand has not moved since I last looked at it, I decide to erase my mind as far as I can, forming a credence in h anew based on the perceptual evidence I have as I am looking at the clock. It's far from obvious that this kind of re-take couldn't provide me with any new evidence relevant for h . Nevertheless, in so far as I am prone to make the same errors on both occasions, it may have less relevance than the evidence I gained the first time round. The case would be very much like one in which I use the same measuring device twice in a row, and boost my confidence slightly as a result of reading the second measurement. If I keep doing this, it is plausible that at each round I acquire less and less evidence, and at some point observing the outcome of a new measurements might simply become evidentially irrelevant.

Now consider a harder case in which, instead of updating in the purely sequential manner described, or going back to re-consider just my original perceptual evidence, I reflect on a more comprehensive body of evidence that includes the perceptual evidence I have at t_0 , but also the evidence about my own doxastic state at t_0 . In re-considering the original perceptual evidence, now as part of a more comprehensive body, an element of the kind of re-take described above may be present. Maybe I could as it were read something off my evaluation of the

⁴⁰ If I always considered my total evidence, learning about the credences I have at t_1 might give me new evidence about p just because they gave me evidence about what the rational way of taking my original evidence $E_{\text{PERCEPTUAL}}$ was. More on this below.

⁴¹ Of course one could, for instance, argue that merely learning that my credences at t_0 were rational could change my estimate of my own reliability, thereby affecting my estimation of what credences I will have at t_1 . But assume that this is not the case: my beliefs about my own reliability are (rationally) robust under evidence about any such single-case successes or failures.

original perceptual evidence from my doxastic response to a more comprehensive body including that evidence. If this is right, then learning about my credence at t_1 might, via this kind of mechanism, give me new evidence about h . However, it is plausible that the resulting boost should not be as great as the initial one – even if I could keep slightly boosting my confidence, I wouldn't be able to get arbitrarily close to 1.

Hence, denying the inertness theses does not lead to the conclusion that by repeatedly learning of my own confidence in some proposition, I could always “bootstrap” myself to virtual certainty that that proposition is true.

Conclusions

I have argued that in so far as evidence about another subject's doxastic states can provide evidence about what one's evidence supports, and about matters such as who will win the next Presidential election, sometime evidence about one's own states can also do so. Evidence about one's own states is neither higher-order nor first-order inert.

Let me conclude with a couple of rather speculative remarks. If what I have said is correct, we may have discovered a new kind of epistemic pathway by which one could find out about the world by finding out about one's beliefs. This pathway that doesn't in any way rely on those beliefs having specific types of externalist content – but merely on the assumption that they are reliable indicators of what the evidence supports, and (hence) reliable indicators of what is true

Does this make for a kind of a priori evidence about matters external to the mind? This depends on whether subjects need some sort of empirical evidence in order to be justified in regarding themselves as reliable evaluators of evidence. But assume that at least sometimes they do not: they are entitled, without any further evidence, to regard their beliefs as fairly accurately reflecting the evidence, and perhaps as fairly good guides to the world. Then, merely learning about my own doxastic states could give me new evidence about the world, evidence that would be at least as a priori as that provided by introspection, or whatever process it is by which I learn about my own mind.⁴²

⁴² I am grateful to Billy Dunaway, Jim Joyce, Brian Hedden, Sophie Horowitz, David Manley, Daniel Morgan, Josh Schechter, Teruji Thomas, Brian Weatherson, and participants at a colloquium at MIT in May 2012.

Appendix

Since we will only be considering credences in one proposition p , and at one time t , I will abbreviate ' $P_0(p) = r_i$ ' as ' O_i ', and ' $\text{Cre}(p) = r_i$ ' as ' M_i '. Also, since the question concerns my credences, I will write ' P ' instead of ' Cre '. With ' \vee ' for disjunction, the assumptions made are then the following:

1. $P(M_i|O_i) > 0.5$
2. $P(\bigvee_{1 \leq i \leq n} O_i) = 1$, and for all i , $0 < P(O_i)$

What I want to show is that for arbitrary $i \in \{1, \dots, n\}$

3. $P(O_i|M_i) > P(O_i)$.

Assume that $i = 1$.

$$4. P(M_1) = P(M_1|O_1) \times P(O_1) + \sum_{j>1} P(M_1|O_j) \times P(O_j)$$

$$5. P(M_1) < P(M_1|O_1) \times P(O_1) + \sum_{j>1} P(M_1|O_1) \times P(O_j) \quad (1., 4.)^{43}$$

$$6. P(M_1|O_1) \times P(O_1) + \sum_{j>1} P(M_1|O_1) \times P(O_j) = P(M_1|O_1) \sum_{j=1}^n P(O_j) = P(M_1|O_1) \quad (2.)$$

$$7. P(M_1) < P(M_1|O_1) \quad (5., 6.)$$

$$8. \frac{P(M_1|O_1)}{P(M_1)} > 1 \quad (7.)$$

$$9. P(O_1|M_1) = \frac{P(M_1|O_1)}{P(M_1)} \times P(O_1) > P(O_1) \quad (8.)$$

Hence, 3. is true.⁴⁴

⁴³ It follows from 1. that for $j \neq 1$ $P(M_1|O_j) < 0.5 < P(M_1|O_1)$

⁴⁴ I am grateful to Teruji Thomas for making this proof much simpler than it would otherwise have been.

Bibliography

Byrne, A.

Christensen, D.

2007 "Epistemic Self-Respect", *Proceedings of the Aristotelian Society* 107: 319-337.
"Ratioanal Reflection"

2010 "Rational Reflection", *Philosophical Perspectives* 24: 121-140

2011 "Disagreement, Question-Begging, and Epistemic Self-Criticism", *Philosophers' Imprint* 11:6.

Dretske, F.

Elga, A.

manuscript: "The puzzle of the unmarked clock and the new rational reflection principle"

Evans, G

Varieties of Reference

Fernandez, J.

2003: "Privileged Access Naturalized", *Philosophical Quarterly* 53:212, pp. 352-372.

Hall, N.

Kelly, T.

Lewis, D.

Van Fraassen, B.

1984: "Belief and the Will", *Journal of Philosophy* 81: 235-256.

White, R.

2009: "On Treating Oneself and Others as Thermometers", *Episteme* Vol 6: 233-250.

Williamson, T.