Philos Stud
DOI 10.1007/s11098-009-9441-3

Rationally self-ascribed anti-expertise

Nicolas Bommarito

Abstract In their paper, "I Can't Believe I'm Stupid," Adam Elga and Andy Egan introduce a notion of anti-expertise and argue that it is never rational to believe oneself to be an anti-expert. I wish to deny the claim that it is never rational for agents like *us* to ascribe anti-expertise to ourselves by describing cases where self-ascribed anti-expertise makes real life agents *more* rational.

Keywords Epistemology · Anti-expertise · Irrationality · Adam elga · Andy egan

1 Rationality and anti-expertise

Egan and Elga (2005) argue that the self-ascription of anti-expertise entails a rational failure, but before laying out their argument it will be helpful to clarify their use of the term 'anti-expertise.' Their use is quite technical; an anti-expert regarding some subject is one who has at least two beliefs about the subject, confidence in at least one, and at least half of the beliefs about this subject are false.[1] Anti-expertise, then, is not simply a lack of confidence in one's beliefs, but rather a confidence that one's beliefs have gone wrong.

This use of the term 'expertise' is somewhat idiosyncratic. When compared with ordinary English usage, this is quite a restricted sense. We think of experts as not

[1] Egan and Elga p. 84. Compare with the use in Sorensen (1987) where it applies only to a specific proposition (see p. 315).

N. Bommarito (✉)
Brown University, Providence, USA
e-mail: nicolas_bommarito@brown.edu

Published online: 13 October 2009

only having true *beliefs*, but also a skill in intuitive *judgment*.[2] Real life experts often appeal to fuzzy heuristics provided by experience—we hear explanations like "When you've looked at as many carburetors as I have, you can just see where the problem is." A skilled logician may "just see" that two complex statements she has never seen before are truth functionally identical, though she did not have a prior belief and does not construct a mental truth-table. These are important aspects that must be included in any account of the notion. However, since my aim is not to give an account of expertise itself, I will limit myself to using Elga and Egan's more technical sense.

Egan and Elga explicitly state two requirements for rationality (I will claim that they implicitly appeal to another, but more on that later):

Transparency: One is correct about one's own beliefs.[3]
Coherence: One's beliefs are (probabilistically) consistent.

As far as rational requirements go, these seem quite plausible. When we ascribe anti-expertise to ourselves, so Egan and Elga argue, we must violate one of the above conditions.

Suppose a friend asks me about my beliefs regarding René Descartes. I mention that I believe he was French, the author of *Meditations on First Philosophy*, and a Rationalist. But then I reflect a bit and notice that it has been nearly 7 years since I learned about early modern philosophers, I am really bad about remembering historical facts, and I have been wrong more often than I have been right when making claims about the subject. So, I caution her that I also believe it likely that I'm an anti-expert regarding Descartes, that my beliefs are most likely wrong about the matter. Assuming my friend does not dismiss my claims of anti-expertise as false modesty, Egan and Elga are right to point out that my statements are likely to have a strange effect on her. She replies, "So you believe *both* that Descartes was French *and* that you are probably wrong about it?" Now I am in trouble; if I believe that I am most likely wrong in these beliefs about Descartes, then it seems that I believe Descartes was *not* French, *not* the author of *Meditations*, *not* a Rationalist. But that contradicts with the first-order beliefs that I reported to my friend—I have violated the coherence condition.[4] Attempting to avoid incoherence I might try to backpedal, "Look, I'm not really sure *what* I believe about Descartes, but I *am* sure it's probably wrong." But by claiming that I lack access to my own beliefs regarding Descartes, I violate the transparency condition. Since I cannot give up coherence or transparency on pain of irrationality, my only option seems to be to give up my claim of anti-expertise. I can do this either by disavowing my original beliefs and so removing the object of my anti-expertise or by giving up the belief in my anti-expertise itself. Either way, my belief in my own anti-expertise must go.

[2] For example, Dreyfus and Dreyfus (2005) claim that expertise is based in unreflective situational responses and intuitive judgments. Also recall the central role Aristotle gives to experience and habituation in the *Nicomachean Ethics*, which he explicitly extends to medicine and navigation.

[3] Egan and Elga claim that their argument works even if this condition is loosened to "decent access to one's own beliefs." (p. 83). Since I do not attack the transparency condition, I leave it in its stronger form.

[4] This is a somewhat informal take on Egan and Elga's conclusion, which involves probabilistic coherence rather than strict logical contradiction.

Egan and Elga suggest that there is something special about the first person perspective that allows us to rationally ascribe anti-expertise to others, but not to ourselves.[5] They describe the case of Professor X, a professor of geology. Though she tries to correct her mistakes, Professor X publishes book after book on geology, each of which turns out to be quite far from the truth. One day she receives a very reliable report about the accuracy of an anonymous geology professor. The report is not good. The professor in question is a considerable anti-expert regarding geology. Egan and Elga claim that at this point, Professor X may reasonably hold her current geological beliefs and also sincerely say to herself, "Wow, the professor in this report sure is an anti-expert about geology." After all, it is no violation of coherence for our beliefs to conflict with the beliefs of *someone else*. Nor does it violate transparency to be incorrect about the beliefs of *someone else*. But, according to Egan and Elga, when Professor X learns that the anonymous professor described in the report is in fact *herself*, she can no longer rationally ascribe anti-expertise.

So what is the rational response for Professor X in this situation? Egan and Elga claim she should simply change:

> When Professor X finds out that the hapless scholar is her, she should become convinced that she has been an anti-expert. But she should do something about it; she should change her opinions about geology in such a way that she is no longer an anti-expert.[6]

In what follows, I wish to push against their advice in these cases as applied to ordinary agents by considering examples where the self-ascribing of anti-expertise makes an agent more rational than they would otherwise be.

2 Rationality for mere mortals

Why do theorists think about ideally rational agents? One answer might be simply that it is fun. It can be intrinsically interesting in the same way comic book collectors may just enjoy speculating about the limits of Superman's powers. Another reply might be that a good way to learn about our concepts and their relationships is by pushing them to their limits. However, I suspect that most uses of ideal agents in philosophy are not merely for these purposes. Rather, ideal agents are invoked because they are supposed to show something about us less-than-ideal agents; after all, Egan and Elga entitled their paper "I Can't Believe I'm Stupid" and not "Ideally Rational Agents Can't Believe They're Stupid."

In suggesting a change of belief (either withholding belief or forming new beliefs) to be *the* rational response, Egan and Elga have added another assumption about rational agents—control. Rational agents are assumed to be able to change their beliefs at will. Perhaps a perfectly rational Professor X would just give up all

[5] This is disputed; Christensen (2007) suggests that this asymmetry follows from more general evidential concerns. In particular, he points to the contingent facts that we typically happen to be in a better position to know about our own beliefs and though our beliefs may line up with those of others, they always line up exactly with themselves. See Christensen pp. 334–335.

[6] Egan and Elga p. 86.

of her geological beliefs, but for the flesh and blood Professor X and the rest of us less-than-perfectly-rational agents, changing our opinions, even about topics we are fairly certain we are mistaken about, can be an arduous task. Consider the following passage from Augustine's *Confessions*, in which he describes his attempt to change his own religious convictions:

> I was saying inside myself: "Now, now, let it be now!" and as I spoke the words I was already beginning to go in the direction I wanted to go. I nearly managed it, but I did not quite manage it. Yet I did not slip right back to the beginning; I was a stage above that, and I stood there to regain my breath. And I tried again and I was very nearly there; I was almost touching it and grasping it, and then I was not there, I was not touching it, I was not grasping it.[7]

It is not hard to picture Professor X, seated at her desk with the report in front of her saying to herself, "Now, now, let it be now!" Like Augustine, her struggle to change her geological beliefs is likely to be a long and difficult task. Is it irrational for her to accept her own anti-expertise during this period? Without this acceptance, how can she continue her difficult conversion?

It can seem that even Perfectly Rational Agents must, even if only for an instant, at some point think some set of their current beliefs to be more likely wrong. After all, isn't that what motivates agents to change their beliefs when they get reliable evidence that they have been mistaken? If you never, at some level, accept that you have gone wrong, why would you change? One might idealize this problem away by imagining that *perfectly* rational agents always respond to evidence instantaneously,[8] but this sort of extreme idealization no longer seems to resemble the flesh and blood agents we seek to understand.

Consider again Egan and Elga's advice for Professor X to withhold. They matter-of-factly explain, "So she should suspend judgment about geological matters."[9] Imagine for a moment that you are the unfortunate geology professor. You have spent the better part of a decade in graduate school learning about geology, enduring long hours studying textbooks and out in the field gathering data. You are tenured faculty at a respected university and have published papers in peer-reviewed journals. And when the secretary slaps this report showing *you* to be a geological anti-expert on your desk, you are supposed to just *give up* all of your hard-earned beliefs in the blink of an eye? Augustine-like difficulties in giving up your hard-earned professional beliefs are to be expected, but you manage to set out on the road by beginning to accept that you are in fact a geological anti-expert. Now on top of this you must deal with epistemologists pointing their fingers and calling you irrational? During this inevitable process of doxastic housecleaning, rather than insisting on the truth of her geological beliefs, it seems like the most rational course for Professor X is for her to form the meta-belief that she is in fact an anti-expert regarding geology.

[7] Augustine, *Confessions* VIII.8 (Rex Warner trans.).

[8] An anonymous reviewer at *Philosophical Studies* suggested that ideal Bayesian agents might escape this problem updating all their beliefs at once by conditionalization, allowing them to be probabilistically coherent at all times.

[9] Egan and Elga p. 87.

3 Sticky irrationality

It is not just temporal constraints that epistemic mortals face; sometimes we have faulty beliefs that no amount of time and cognitive willpower can eliminate. Cory and Daniel both find themselves with a set of beliefs regarding aliens; they accept that aliens exist, visit Earth, scan the brains of humans, and occasionally abduct them during sleep and implant tracking devices. Neither can manage to shake these beliefs no matter what they try. Daniel treats his belief as a bit of sticky irrationality resulting from certain traumatic experiences, an imaginative childhood, and a wish for paranormal coloring of his rather drab day-to-day life.[10] He concludes that he must be wrong about alien visitors and so believes himself to be an anti-expert on the subject. But Cory, having just finished reading Egan and Elga's paper, decides that since the Perfectly Rational Agent never ascribes anti-expertise to herself, neither should he. If ascribing anti-expertise involves some rational failure, Cory wants no part of it. He laughs derisively when Daniel must utter somewhat Moore-paradoxical sentences like "I have beliefs about aliens visiting Earth, but I believe most of these are false." Cory concludes that the rational meta-belief to form is that he is in fact correct about the alien visitors.

Both Daniel and Cory are irrational; their beliefs regarding aliens do not respond appropriately to the evidence. Both have made at least one mistake, but in denying his anti-expertise, Cory makes another. He may rest easier knowing that at least *he* doesn't have to assert incoherent Moore paradoxical claims, but Daniel at least gets it right about his mistake, even if he cannot do anything to correct it. Daniel's meta-belief that he is an anti-expert *does* respond to the evidence properly, but at the cost of coherence. Cory, in deciding to take his beliefs as likely to be true, manages to preserve coherence but also allows his initial irrationality to spread. If Cory *really* does not accept his anti-expertise, *he* will be the one covering his head with aluminum foil to prevent brain scans and cutting up his arm to remove the tracking device, all the while taking himself to be totally rational in doing so. And even if Daniel should find himself with aluminum foil on his head too, at least he will be sensible enough to consider himself misguided in doing so.

The story of Cory and Daniel might strike one as an extreme case, but less dramatic examples of this kind of phenomenon are quite common. It is a psychological fact that the pull of some fallacies disappears when revealed to us, while others do not. Sitting in the seminar room I may reason, "Derek is absent and if someone is sick they will be absent, so Derek must be sick." When another more analytically sharp student points out that I have just affirmed the consequent, the conclusion seems much less cogent to me, even on a gut level. However, at the casino I may witness a string of nine consecutive red outcomes at the Roulette

[10] In her *Abducted*, Clancy (2005) gives a detailed account of how otherwise rational people come to believe they have been abducted by aliens. Most are aware of other more pedestrian explanations and reject them precisely because they feel such explanations do not fit the data as well. For example, an appeal to 'sleep paralysis' fails to capture the intensity of the distressing emotions and mysterious quality of the experiences. Paranormal explanations are also able to give meaning and significance to experiences in ways that more clinical and scientific explanations cannot; recall that Carl Jung referred to extraterrestrials as 'technological angels.'

wheel and conclude, "Black *has to be* due!" This time, when my friend reminds me that the outcomes are independent and I have just succumbed to the Gambler's Fallacy, I understand my mistake, but it does not remove the pull of believing that black *really is due*.[11] Sticky irrationality of this kind is all too common in the lives of epistemic mortals like us. Eoin may know full well that his watching the rugby match between Australia and Ireland can have no effect on the outcome, but still feel that if he watches, Ireland will lose. Nordrun may be quite confident that she can pass her TOEFL exam, but still feel that if she says such a thing out loud, she will jinx it. The rational response for Nordrun and Eoin does not seem to be insist on the rationality of their respective beliefs, but rather to accept them as slightly endearing and perhaps even psychologically helpful irrationalities.

4 I *can* believe I'm stupid

Whether or not Perfectly Rational Agents ever find themselves self-ascribing anti-expertise, such self-ascriptions can often be the most rational choice for epistemic mortals like us. In the fortunate cases where we find ourselves able to consciously change our beliefs or withhold regarding a topic, such doxastic changes often require a good deal of time to execute. During this time, to deny our own epistemic failures is not only to be in dishonest with ourselves, but also to rob ourselves of one of the strongest motives to keep attempting to bring about a change.

In cases of more sticky irrationality, failure to acknowledge our own mistakes may make us more coherent, but at the expense of other epistemic virtues like accuracy and humility. Such an attachment to coherence allows mistakes to spread and, like Cory, can make an agent much *less* rational. As imperfect agents, we are bound to get things wrong and to reason poorly, but responding to these failings by denying them eliminates any hope of correcting or containing them. Being in denial of our own rational and epistemic failings does not make us more rational; those who do not accept their mistakes not only make another, but fail to learn from the first. Not only can we believe that we are stupid, but when we have good evidence of our own stupidity, we should believe so.[12]

References

Christensen, D. (2007). Epistemic self-respect. In *Proceedings of the Aristotelian Society* vol. CVII part 3.
Clancy, S. (2005). *Abducted.* Cambridge: Harvard University Press.
Dreyfus, H. L., & Dreyfus, S. E. (2005). Peripheral vision: Expertise in real world contexts. *Organization Studies, 26*, 779.

[11] My conclusions in these cases need not be a single belief, but may result in a set of beliefs about which I can have Egan and Elga style anti-expertise; I may believe Derek is sick, he is at home, he would like chicken soup, etc. or black is due, I should bet on black, I should bet a lot, etc.

[12] Thank you to David Christensen and Eoin Ryan for reviewing drafts and discussing these issues with me (often at length!) and also to an anonymous reviewer from *Philosophical Studies* for helpful comments.

Egan, A., & Elga, A. (2005). I can't believe I'm stupid. *Philosophical Perspectives, 19*, 77–93.
Sorensen, R. A. (1987). Anti-expertise, instability, and rational choice. *Australasian Journal of Philosophy, 65*(3), 302–315.