Akrasia and Traitors

Anonymous

2025-04-10

Bar Luzon argues that akrasia is irrational because it leads to violating a principle called **Avoid Treachery**. In response, I argue that Avoid Treachery is insufficiently motivated, that it presupposes a picture of rational inference that defenders of akrasia have independent reason to reject, and that there are models in which Avoid Treachery is false.

The Debate

A prominent debate in recent epistemology has been whether it can ever be rational to believe propositions of the form of **SA**, or of some similar forms.

SA q and it is irrational for me to believe q.

The *enkratic* philosopher says all beliefs of that form are irrational. The *anti-enkratic* philosopher says that they are sometimes rational.

The debate here isn't always about **SA** (for Simple Akrasia), it could also be about **LA** (for Likely Akrasia).

LA q and it is probably irrational for me to believe q.

In this note I'm mostly not going to distinguish these, or several other nearby forms, but the differences between them will lurk in the background. This is in part because Timothy Williamson (2014) agrees with the enkratic philosophers about **SA**, but the anti-enkratic philosophers about **LA**, and because Williamson's ideas loom so large over the debate.

Recently in this journal, Bar Luzon (Forthcoming) has argued on the side of the enkratic philosophers. She argues that the anti-enkratic philosophers say it is rational to violate the principle she calls **Avoid Treachery**.

Avoid Treachery (AT) For every proposition p and for every positive epistemic status E, if one knows that [p has E for one only if p is false], then one ought not believe p.

In this principle, E ranges over the statuses epistemic justification, epistemic rationality, evidential support and epistemic permissibility, and the conditional is a material conditional. The 'ought' is purely epistemic; if one thought belief in God was justified on Pascalian grounds one wouldn't be moved by an argument from **AT**. So I'll take 'one ought not believe p' to just be that it's (epistemically) irrational to believe that p. So we can formalise **AT** as follows, where KA is that hero knows A, for one of the four statuses at the start of this paragraph, EA is that p has that status for hero, and RA is that hero rationally believes A.

AT Formalised $K\neg(p \wedge Ep) \rightarrow \neg Rp$

The argument for the irrationality of **SA** follows pretty quickly. Let p be $(q \wedge \neg Rp)$. Assume $E(A \wedge B)$ implies EA, that RA implies EA, and that hero knows anything that can be proven in a few lines of logic. Then it's easy to show $K\neg(p \wedge Ep)$, and hence $\neg Rp$, which just is the enkratic view.

The point of this note is to argue that the anti-enkratic philosopher has good reasons to reject **AT**. It would be too easy to say that their view implies it is false. Of course it does, since Luzon's argument against their view is valid! What I want to argue is that the reasons behind the anti-enkratic view give us somewhat independent reason to reject **AT**. I'm going to offer the following arguments against **AT** in sections **?@sec-math** to **?@sec-model**.

1. **AT** fails for other nearby values of E, and this undermines the motivation for believing it holds for these values.
2. The argument for **AT** turns on an idea about how evidence guides belief that the anti-enkratic philosopher rejects.
3. There are plausible models for evidence and belief where **AT** fails.

But first it helps to rehearse the arguments for the anti-enkratic view, to see how these objections flow from them.

The Arguments

Simplifying greatly, the anti-enkratic position relies on one presupposition, followed by one of two (independent) arguments. The presupposition is easiest to see with an example.

Hero has a faculty meeting today, but they have forgotten about it. Fortunately, they just got a reminder email from their chair saying there's a meeting today. Now they believe, indeed know, there's a meeting today.

The presupposition is that the following three things are in principle distinct.

1. Hero's reason for believing that there is a meeting today, i.e., the email they got from the chair.

2. The facts that make the email from the chair a reason to believe there is a meeting today. Just what those are turns on the full theory of testimony, but presumably they include things like the chair's reliability, the frequency of emails being faked, and so on.

3. The reasons hero has for believing that the email is a reason to believe there's a meeting today.

The presupposition is that 1 and 3 are distinct. The reason that they are distinct is that 1 and 2 are distinct, and 3 requires hero to have thoughts about (or at least sensitive to) 2, while 1 does not.

With that in place, the first argument for the anti-enkrasia view starts with anti-

exceptionalism about epistemology.[1]

Luzon, Bar. Forthcoming. "Epistemic Akrasia and Treacherous Propositions." *Philosophical Quarterly*, Forthcoming.

Williamson, Timothy. 2000. *Knowledge and its Limits*. Oxford University Press.

———. 2011. "Improbable Knowing." In *Evidentialism and Its Discontents*, edited by Trent Dougherty, 147–64. Oxford: Oxford University Press.

———. 2014. "Very Improbable Knowing." *Erkenntnis* 79 (5): 971–99. https://doi.org/10.1007/s10670-013-9590-9.

---

[1]For anti-exceptionalism about logic, see (?)