# Epistemic Akrasia and Treacherous Propositions

Bar Luzon, Rutgers University

Forthcoming in *Philosophical Quarterly*

**Abstract:** I argue that one ought not be epistemically akratic. Although this position may look self-evident, it is hard to pin down exactly what's wrong with the akratic subject. Indeed, some philosophers argue that epistemic akrasia is permissible. The standard anti-akratic response focuses on the weird *downstream* implications of this state for action and assertion. This approach, however, is unsatisfactory, since it fails to explain the *epistemic* impermissibility of epistemic akrasia. Here, I argue that epistemic akrasia is impermissible on a purely epistemic basis. Doing so brings out the connection between positive epistemic statuses and truth: it is that connection that explains the fundamental impermissibility of epistemic akrasia.

Consider the following case:

> **UNDERGRADUATE SKEPTIC**
> Luna is just starting out in philosophy. Her epistemology professor introduces her to skeptical arguments that are, so far as she can tell, flawless. Furthermore, Luna's professor presents them as if they were in fact flawless, highlighting many failed attempts to rebut them and no successful ones. As a result, Luna forms the belief that she shouldn't believe that she has hands. Still, Luna cannot stop believing that she really has hands; she keeps seeing them everywhere![1]

In this case, Luna is epistemically akratic: for the proposition $p$ = *Luna has hands*, she believes both $p$ and that she shouldn't believe $p$. More generally, a subject is epistemically akratic whenever they believe both $p$ and $p$ *lacks* $E$ *for me*, where $E$ is any of the following all-things-considered, non-factive positive epistemic statuses: epistemic justification, epistemic rationality, evidential support and epistemic permissibility.[2] In this paper, I argue that one ought never be epistemically akratic. Regardless of how good your evidence for both $p$ and $p$ *lacks* $E$ *for you* may seem, you are never epistemically permitted to believe both at once.[3,4]

---

[1] This case is taken from Greco (2014), p. 204.

[2] When I say that a proposition has some positive status, I mean it in the sense where it requires actually having an epistemically valuable feature, not just that if I were to believe in that proposition, it would have some epistemically valuable feature.

[3] By 'epistemically impermissible', I mean to denote the all-things-considered *epistemic* ought—it's supposed to capture the ultimate verdict of the right epistemic norms. I use 'ought' and 'permissible' as duals, and I use 'impermissible', 'ought not' and 'shouldn't' interchangeably. All the other relevant epistemic statuses that a belief can lack are plausibly closely related to this ultimate verdict. When it's relevant for the argument of the paper and to the extent that it is, I take a stand on some relations that hold between the other ways of specifying $E$ (justification, rationality, evidential support and epistemic permissibility).

[4] Going forward, I often drop the 'epistemic' in front of the relevant statuses.

This position may look self-evident. After all, epistemic akrasia certainly *seems* problematic. Epistemically akratic subjects take themselves to be failing by their own lights, and so betray some kind of incoherence.[5] Moreover, from the akratic subject's own perspective, this situation should seem epistemically suspect.[6] For it's unclear what such subjects should think about their own epistemic situations—should Luna, for example, think that she is oh-so-lucky to believe the truth despite lacking good reasons for believing it?

Still, it is hard to pin down exactly what's wrong with the akratic subject. Indeed, some philosophers argue that epistemic akrasia (in one form or another) is permissible.[7] After all, it seems like you can have good evidence for $p$ while having good, though misleading, evidence telling you that you lack evidence for $p$.[8] If one is permitted to believe in accordance with one's evidence, then it seems like you should be permitted to believe akratically.

The standard anti-akratic response to these arguments focuses on the weird *downstream* implications of this state for action and assertion.[9] For example, Sophie Horowitz (2014) points out that subjects like Luna could reason into thinking that it's a good thing they haven't followed the evidence where it leads, for otherwise they would end up with a false belief (in Luna's case, about whether she has hands). This approach, however, is unsatisfactory, since it fails to explain the *epistemic* impermissibility of epistemic akrasia. There is something wrong with epistemic akrasia from a purely epistemic vantage point, regardless of norms that link permissible beliefs to action or assertion.

By appealing to a 'purely epistemic vantage point', I'm not intending to legislate how people should use the term 'epistemic'. Rather, I'm drawing on a distinction that should by now be familiar. Take, for example, arguments for probabilism about credences.[10] Philosophers working on this issue have distinguished between Dutch book arguments, which focus on the weird betting behavior of subjects whose credences don't obey the probability calculus, and accuracy arguments, which focus on showing that non-probabilistic credences aren't accurate in some important sense.[11] While both arguments identify bad features of not aligning one's credences with the probability calculus, the former captures bad downstream consequences of such misalignment, while the latter aims to capture what's epistemically or fundamentally wrong with having such credences. And presumably, the latter bad feature partly explains why such credences exhibit these bad downstream consequences.

---

[5] Some philosophers take epistemic akrasia to be impermissible due to its purported incoherence. See, for example, Titelbaum (2015).

[6] Horowitz (2014) develops this line of criticism.

[7] See Coates (2012), Williamson (2014), Christensen (2016), Weatherson (2019) and Lasonen-Aarnio (2020).

[8] Or, in Christensen (2016)'s case, cases in which the subject has good evidence for $p$ as well as good evidence for their anti-reliability in assessing her first-order evidence.

[9] See, for example, Horowitz (2014) and Littlejohn (2018). Greco (2014) advances a completely different line of response that has to do with meta-epistemic commitments.

[10] I thank Kevin Dorst for proposing this analogy to me.

[11] See Joyce (1998).

Similarly, subjects with straight-out inconsistent beliefs would undoubtedly reason and act in weird ways. However, it's implausible to think that this is what is fundamentally wrong with them. The state they're in—that of believing both *p* and not *p*, for some *p*—is itself epistemically problematic. Even if the subject doesn't go on to assert anything or act on the basis of this state in any way, there is still something wrong with the subject being in the state. And again, if this state does have weird downstream consequences for the subject, it's at least partly due to it being epistemically problematic that it has these consequences.

In what follows, I argue that epistemic akrasia is impermissible on a purely epistemic basis, in the sense elucidated above. This reveals what is wrong with the state of akrasia itself, rather than with its downstream consequences. Doing so also brings out the connection between positive epistemic statuses and truth: it is that connection, I argue, that explains the fundamental impermissibility of epistemic akrasia. Hence, reflecting on the epistemic status of epistemic akrasia helps us get a better understanding of the very nature of important epistemic statuses such as justification and evidential support.

To give the full argument, more work needs to be done. But the gist of the argument can be given right away. Assume I'm epistemically akratic, and that for some proposition *p*, I believe the conjunction: *p, but p lacks E for me*, where *E* stands for permissibility, rationality, justification and evidential support. Here, for concreteness, let *E* be evidential support. So I believe the conjunction *p, but p lacks evidential support for me*. Then I have a belief that cannot be both true and evidentially supported at the same time. If my belief is true, then the first conjunct isn't evidentially supported (that's what the second conjunct says!), and hence the whole conjunction isn't. And if my belief is evidentially supported, then the second conjunct is false, since a conjunction can be evidentially supported only when both of its conjuncts are evidentially supported. Hence the whole conjunction is false. Now that I have followed this argument, I can easily know that my belief cannot be both true and evidentially supported. And I ought not hold any belief that I can easily know cannot be both true and evidentially supported.

In order to turn this gist into a full, convincing argument, we need to do a few things. First, the argument relies on the claim that we ought not believe those propositions that we know cannot be true and evidentially supported at the same time. This claim requires both elucidation and argumentation. Second, we need to explain how we can move from the impermissibility of believing the conjunction to the impermissibility of believing both conjuncts, which is what we defined epistemic akrasia as. And finally, we need to show that all of this applies not just in the case where *E* = evidential support, but also where *E* = rationality, justification or permissibility. This is the toll of the rest of the paper.

The plan for the paper is as follows. §1 introduces and motivates the key principle on which the argument relies. §2 presents the argument in a more precise form, as well as a second argument for the transition from the impermissibility of believing the conjunction to the

impermissibility of believing its conjuncts. §3 situates the discussion in §§1-2 in the literature about epistemic akrasia. §4 concludes with an analogy to Moorean propositions.

## §1 Avoiding Knowable Treachery

For the rest of the paper, let $E$ range over the positive epistemic statuses of justification, rationality, evidential support or permissibility. Now, consider the following principle:

> **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 $p$.

where all the conditionals in AT are material conditionals.[12] Call a proposition $p$ treacherous in a given case if and only if, in that case, [$p$ has $E$ for one only if $p$ is false].

To see how AT works, consider an epistemic nihilist who doubts that there are any truths about epistemic rationality and epistemic permissibility. Her cousin, the moral nihilist, believes that any positive claim about morality—for example, that torturing your cat is wrong—is false. Similarly, the epistemic nihilist believes that any positive claim about epistemic permissibility—for example, that one is permitted to believe that it is raining—is false. Let $p$ be the proposition that, for any proposition $q$, it is not true that one is permitted to believe $q$. Clearly, one is permitted to believe $p$ only if $p$ is false. If $p$ was true, one would be permitted to believe nothing. Since this is so clear, you now know it. And since you know that [you are permitted to believe $p$ only if $p$ is false], AT tells you not to believe $p$.[13]

Why accept AT? A belief in a proposition $p$ that satisfies the antecedent of AT is self-undermining: whatever positive epistemic status $E$ it presumably has, that status cannot be had together with truth. That is, whenever the belief has that positive epistemic status, it is guaranteed not to be true. Why do we care about positive epistemic statuses such as rationality or evidential support? Presumably, these statuses are important insofar as they guide us toward the truth. We care about forming accurate representations of the world, and we're more likely to do so if our beliefs are justified, rational, supported by our evidence, etc. In the case of treacherous propositions, there's a sense in which the relevant epistemic status $E$ malfunctions: instead of helping the subject reach an accurate conclusion about $p$, it holds only when $p$ isn't the case. Moreover, if the antecedent of AT applies in a given case, the subject also *knows* that epistemic status $E$ malfunctions in that case. Given that knowledge provides both factivity and some robust epistemic status, then the subject also has a factive, robust awareness of this self-undermining.

---

[12] Variants of AT might be plausible for other conditionals (e.g. the indicative, the strict), but addressing those possibilities would take us too far afield.

[13] This argument is not going to convince the epistemic nihilist, who presumably also doubts AT, or does not care about whether one 'ought' to believe something or not. This is just an example of how AT works, not a decisive argument against epistemic nihilism. I thank Patrick Wu for suggesting this analogy to me.

Now, although epistemic statuses aren't infallible—a belief can be rational, justified, etc. and yet false[14]—there's still an important sense in which they *cannot* malfunction if we take seriously the idea that epistemic statuses are guides to truth.[15] To take this idea seriously, I suggest, one must posit a strong connection between truth and positive epistemic statuses.[16] Now, this connection is notoriously hard to pinpoint, and I don't have a general theory of how truth and positive epistemic statuses interact, but that there is some such connection should be uncontroversial. Moreover, such a connection should satisfy, at the very least, the following principle:

> **Avoid Falsehoods (AF)**
> For every proposition $p$ and for every positive epistemic status $E$, if one knows that $p$ is false, then $p$ lacks $E$ for one.

Once I know that $p$ is false, it directly affects any other positive epistemic status a belief in $p$ may have. If AF wasn't the case, epistemic statuses couldn't be good guides to truth. They wouldn't only be fallible, but would also be insensitive to truth in a much more concerning way. It's the same kind of insensitivity to truth that we would attribute to a subject if they believed some proposition albeit knowing that it's false. The same way we would consider the latter to be a bad epistemic agent—in this instance for not being attuned to the truth—we should consider epistemic guides which point to $p$ when the subject knows that $p$ is false to be bad epistemic guides. And indeed, most theories of evidential support, for example, would take my knowledge that $p$ is false to be *decisive evidence* against $p$, and so if I know that $p$ is false, $p$ would lack evidential support for me.

I think that the same sensitivity towards the connection between positive epistemic statuses and truth that leads us to accept AF should lead us to accept the following principle:

> **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 $p$ lacks $E$ for one.

If a belief in a known treacherous proposition could have had status $E$, then $E$ could not have been a good guide to the truth, and, *ipso facto*, couldn't have been an all things considered

---

[14] Pace Littlejohn (2012).

[15] This idea has been taken up even by philosophers who find it hard to reconcile positive epistemic statuses (such as rationality and evidential support) and truth. For example, albeit arguing that from a certain perspective, conciliationism about higher-order evidence doesn't maximize expected accuracy, Schoenfield (2015) suggests a different picture of rationality (one that ties is to the best plan to make, rather than to the best plan to follow) on which being rational, according to the conciliationist, does maximize expect accuracy after all. Similarly, Horowitz (2019), after arguing that higher-order evidence is predictably misleading, claims that no particular piece of higher-order evidence can be recognized as misleading from the subject's point of view. This, I think, is an indication that the plausibility of accounts of positive epistemic statuses largely depends on their ability to deliver this important connection to truth. Finally, Williamson (2000) uses this connection between truth and justification to argue for E=K, the claim that "knowledge, and only knowledge, constitutes evidence" (Williamson 2000, p. 185).

[16] See Dorst (2020) for an account of evidential support that secures such a strong connection.

positive epistemic status. To see the appeal of AT*, consider an analogy with testimony. Tom witnessed the murder of Myrtle, which is now investigated by Harry. Imagine that Tom is a bad witness in the following sense: he knows full well that if he says anything about the murder to Harry, it will be false. If Tom ends up telling Harry anything about the murder, he is a bad witness, since he acts as a bad guide to the truth. Similarly, if a positive epistemic status $E$ "knew" some proposition $p$ to be treacherous and yet told the subject to believe in $p$, $E$ would have been a bad guide to the truth. The sense in which $E$ "knows" that it can only recommend a belief in a falsehood is this: plausibly, if one knows that $p$, then $p$ has $E$ for one; so from the fact that the subject knows that [p has $E$ for one only if $p$ is false], it follows that the subject's belief that [p has $E$ for one only if $p$ is false] has $E$. Again, all things considered epistemic statuses may misfire and point me away from the truth: my evidence may support a belief in a falsehood, I can be justified in believing a falsehood, etc. But these statuses can't be so detached from the truth that they point me away from the truth while "knowing full well" that that's what they're doing. Just as Tom is a bad epistemic guide, so would be a status $E$ which violates AT*.

So we should accept AT* because it's a natural extrapolation from AF. And we can easily infer AT from AT*. To make this easier for the reader, let me repeat the two principles here, differences bolded.

> **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 $p$ lacks $E$ for one.**
>
> **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 $p$ (i.e. $p$ is impermissible for one).**

When $E$ = permissibility, AT* implies AT directly. But even when $E$ is one of the other epistemic statuses, AT seems to follow quite naturally. Consider the example of $E$ = evidential support. *P* lacking evidential support isn't the same as one being forbidden to believe that $p$, but there's quite a tight connection between the two. Presumably, the former implies the latter: if one's evidence doesn't support $p$, it follows that one ought not believe that $p$.

This point shouldn't be controversial. It is implied by the standard evidentialist thought (one ought to believe $p$ iff $p$ is supported by one's evidence), a thought usually accepted by proponents of the permissibility of epistemic akrasia.[17] But it is also reasonably weaker than that thought, since, for example, pragmatic encroachers could accept it.[18] The same kind of line applies to justification and rationality: however one understands those notions, it seems

[17] For example, the model in Williamson (2014) in defense of epistemic akrasia is evidentialist (at least modulo the choice of a prior probability function).
[18] This condition is compatible, for example, with Fantl and McGrath (2002).

that a belief having them is at least a necessary condition for that belief being permissible. So AT* implies AT.

Notice that the case just considered is importantly different from a case where the subject *merely believes* (and even justifiably so) that one's evidence doesn't support *p*. The factivity of knowledge helps us avoid begging the question against the pro-akratic. I'm not claiming that *merely believing* that one's evidence doesn't support some proposition implies that one ought not believe that proposition—that's part of what's at issue between the pro- and the anti-akratic! Rather, I claim that a proposition *in fact not being supported* implies that one ought not believe it.

The reasoning so far might be taken to presuppose that accuracy is the aim of belief, but a similar point can be made if we take knowledge to be the aim of belief instead. Presumably, a belief having *E* is a necessary condition for knowledge. (Admittedly, some externalists may think that this is more plausible with evidential support than with justification, to the extent that the latter differs from the former, but it's still pretty plausible for all of these statuses.) If this is true, then a proposition that satisfies the antecedent of AT is one that cannot be known. This is because a belief in any such proposition cannot both have *E* and be true simultaneously. If we're permitted to believe only those propositions that we can know, then we're not permitted to believe any proposition that satisfies the antecedent of AT.

Assume with me, then, that AT is true, and consider the following case: Ron doesn't know [if Ron's belief in *p* has *E*, then *p* is false], but this is only because he hasn't thought enough about this conditional to form a belief either way. Still, Ron does have enough evidence to be in a position to know that the conditional is true (for concreteness, we could let *p* be our previous claim of epistemic nihilism). Intuitively, Ron is not off the hook—he cannot permissibly believe *p*. So we can strengthen our principle to:

> **Avoid Knowable Treachery (AKT)**
> For every proposition *p* and for every positive epistemic status *E*, if one can easily know that [*p* has *E* for one only if *p* is false], then one ought not believe *p*.

A subject can easily come to know some (true) proposition when that proposition is highly *accessible* to the subject given the subject's evidence, when that proposition is in some sense at the reach of the relevant subject, even if the subject doesn't in fact reach it. One way to gloss the notion is as follows: a subject can easily come to know some proposition *p* when there is an easy transition from the subject's evidence to that proposition. Admittedly, the notion of 'can easily come to know' is vague and context-sensitive, and it can apply to different propositions for different people at different times. Nonetheless, I take it that there are clear cases in which one can easily come to know a proposition. If I know that Albus is wearing a purple cloak, then I can easily come to know that Albus is wearing a cloak. If a father who watches soccer regularly knows that his toddler daughter is playing soccer, he can easily come

to know that she is likely to get hurt, given what he knows of the risks involved in soccer and children's susceptibility to harm.[19]

My use of 'can easily come to know' will be fairly restrictive: I will not assume that one can easily come to know any true proposition that is sufficiently supported by one's evidence and that one is psychologically capable of knowing.[20] Relatedly, it is not the case that one can easily come to know anything that follows from one's evidence regardless of how difficult it is to see the derivation. Only derivations that are, intuitively, fairly simple or obvious will count. Finally, 'can easily come to know' here is implied by 'knows'—if a subject knows that their evidence is treacherous with respect to *p*, then it's also true that the subject can easily come to know that their evidence is treacherous.

## §2 The Impermissibility of Epistemic Akrasia

With this principle in hand, we can now run the full version of my argument for the impermissibility of epistemic akrasia. Start by defining, for an arbitrary but fixed *p*, and for each positive epistemic status *E*:

**Akratic Conjunction (AC) :=** The proposition *q, but q lacks E for one.*

With AKT in hand, we can easily argue that AC ought not be believed:

(1) If one can easily come to know that [AC has *E* for one only if AC is false], then one ought not believe AC. (AKT applied to *p* = AC.)
(2) One can easily come to know that [AC has *E* for one only if AC is false].
(3) Thus, one ought not believe AC. (By MP from 1,2)

I have already defended AKT, and so premise 1. I turn now to premise 2. The reason to accept premise 2 is that there is a simple derivation (fleshed out below) of the fact that AC is treacherous from an extremely plausible principle. The argument shows that AC is treacherous by its very nature: the very thing that AC says makes it treacherous. This means that AC is treacherous not only if we take the conditional in the principle to be the material conditional, but also if we take it to be strict. That is, it's not only that AC having *E* implies its falsity—it's also necessarily so. Since the argument provided below is so simple, its conclusion is not only true, but also easily knowable by any subject. We may even want to say that, regardless of whether one thinks that knowledge always (or usually) implies rational certainty, the argument for AC being treacherous gives us also the latter, rather than just the former.[21]

---

[19] This example comes from Le Morvan (2012) via Goldberg (2017).
[20] See Williamson (2000) on 'being in a position to know'.
[21] If it's easily knowable that AC is treacherous, why hasn't it been noticed by other philosophers that (1) AC is treacherous, and (2) that we can easily come to know this? How can it be, that is, that something that we can easily come to know—the fact that the akratic conjunction is treacherous—is so hard for many of us to

The argument for AC being treacherous goes as follows:

(4) **Epistemic Distribution over Conjunction (EDC):** For all propositions *r* and *s* and for every positive epistemic status *E*, if *r and s* has *E* for one, then *r* has *E* for one and *s* has *E* for one.
(5) Thus, if AC has *E* for one, then *q* has *E* for one. (From 4, EDC applied to *r and s* = AC.)
(6) Thus, if AC has *E* for one, AC is false. (From 5, since *q* having *E* conflicts with the second conjunct of AC.)

The only independent premise in our argument is premise 4, so let me briefly defend EDC. Think of having any of the positive epistemic statuses we considered as a way for a belief to be epistemically okay. Plausibly, whenever it is epistemically okay to believe a certain conjunction, it is also epistemically okay to believe each of the conjuncts. For concreteness, plug 'rational' in for *E*. Plausibly, whenever one is rational in believing a conjunction, one is also rational in believing the conjuncts. For example, say I read in a reliable history book (the reliability of which I'm aware) that Henry VIII had six wives and that the last one outlived him. It seems like if it's rational for me to believe that *Henry VIII had six wives and the last one outlived him*, it is also rational for me to believe each of the conjuncts, namely, that *Henry VIII had six wives* and that *Henry the VIII's last wife outlived him*. And the same goes for any of the other candidates for *E*.

Notice further that since the conclusion of my argument is that epistemic akrasia is always impermissible, we need not assume EDC in full generality. Instead, it's enough to assume that if epistemic akrasia is permissible, then it is possible that epistemic akrasia is permissible for a subject satisfying EDC. And presumably, all cases in which it's even tempting to think that epistemic akrasia is permissible are ones in which both conjuncts of AC have *E*. In fact, proponents of epistemic akrasia never give us a reason to believe the conjunction directly; rather, they provide reasons to believe each of the conjuncts. If EDC doesn't hold in full generality, then, the argument should still go through.[22]

Some might insist that, even though AC cannot be permissibly believed, epistemic akrasia might still be permissible. For while my argument shows that one ought not believe that AC, it doesn't follow that one ought not believe *p* and believe that *p lacks E for one* simultaneously, which is how we characterized epistemic akrasia.

---

recognize? I'm not sure I know the answer to this question, but my guess is this. Most proponents of epistemic akrasia rarely think of akratic conjunctions as such, but rather focus on thinking of subjects who hold two distinct beliefs in the conjuncts. And even philosophers who think of epistemic akrasia as mostly impermissible (such as Horowitz 2014) focus on the subject's attitudes towards *p*—the proposition towards which the subject has akratic attitudes—rather than on the akratic conjunction itself. And attending to the akratic conjunction itself is a key stage in reaching the conclusion that we ought not be akratic—it's only when we think about akratic propositions that we see what accepting the permissibility of epistemic akrasia commits us to.

[22] EDC doesn't hold for every epistemic status. For example, as a referee for another journal has helpfully pointed out, a belief in a conjunction can be sensitive even if one of the conjuncts isn't. Consider the conjunction: I have hands, and I am not a handless brain in a vat.

Before answering this objection, let me first note this: even if the reader ends up rejecting everything I say from now on, there's still an important sense in which the reader agrees that something like epistemic akrasia is epistemically impermissible. That is, the reader agrees that one ought not believe the akratic conjunction for any *p*. And that's already a serious concession on behalf of the pro-akratic.

For my argument to work as an argument for the impermissibility of epistemic akrasia in the way we originally characterized it, the following needs to be true:

> **Conjuncts**
> If one ought not believe that AC, one ought not believe its conjuncts either.

Conjunct is clearly a very intuitive claim. If Conjuncts is false, then a subject who believes the conjuncts *q* and *q lacks E for one* must be doing better than a subject who believes the conjunction *q and q lacks E for one*. Would that be acceptable in nearby cases? Should we think, for example, that although believing *q and not q* is problematic, there's nothing wrong with believing both *q* and *not q* simultaneously without forming the conjunctive belief? To rehearse a familiar point in the context of knowledge, it would be very weird if one couldn't generally extend one's permitted beliefs using deductive inference.[23]

But the reader may worry that Conjuncts cannot be true, because of well-known difficulties with similar principles. Consider the lottery paradox. I'm participating in a 1000-ticket lottery which I know to be fair and to have one (and only one) winner. It seems permissible for me to believe, for each ticket, that that ticket wouldn't win, as it has such a low probability of being the winning ticket. I'm also permitted to believe that some ticket would win, given the setup. But it seems like it's impermissible for me to believe the conjunction of all of those propositions. That is, it's impermissible for me to believe that *ticket 1 wouldn't win and ticket 2 wouldn't win and… ticket 1000 wouldn't win and some ticket would win*. This is because this proposition is inconsistent given what I know about the setup, and that I am aware of this inconsistency. We can avoid the paradox if it is sometimes permissible not to believe a conjunction even while believing the conjuncts. Or so the thought goes.[24]

Using this discussion as a line of objection to Conjuncts faces a problem. Setting aside the issue of whether this is a good solution to the lottery paradox,[25] we should notice the following important feature of lottery cases: in such cases, one's rational credence in each conjunct is extremely high, but one's rational credence in the conjunction is extremely low. In the case of epistemic akrasia, this needn't be the case, and many proponents of epistemic akrasia are committed to it not being the case. The main arguments for the permissibility of epistemic akrasia, for example from Williamson (2011), entail that the degree to which the conjuncts of AC can be evidentially supported is arbitrarily high,[26] in which case the degree to which AC

---

[23] See e.g. Williamson (2000), p. 117.
[24] For an example of such a strategy, see Foley (1979).
[25] For criticisms, see Ryan (1991) and Nelkin (2000).
[26] See Williamson (2011, 2014).

itself can be evidentially supported is itself arbitrarily high. If the better the support for the conjuncts, the better the case for the permissibility of epistemic akrasia, then Conjuncts is true exactly when the permissibility of epistemic akrasia seems most convincing.[27] More generally, let me repeat a point I made in the context of discussing EDC: for the purposes of my argument, all we need is the truth of the following conditional: *if* epistemic akrasia is permissible, then it is possible that epistemic akrasia is permissible for a subject satisfying Conjuncts. And this is guaranteed to be true by Williamson's argument for epistemic akrasia.

## §3 Situating in the Literature

Although this paper's aim is to provide a new argument in the literature about epistemic akrasia, rather than comment on existing ones, a reader familiar with this literature may reasonably wonder how the considerations presented in this paper interact with this literature. Where would pro-akratic philosophers get off the boat, for example? So let me say a little about how this paper fits into the broader literature.

Basically, I think that the motivation for AT* (and similarly AT) is what lies at the heart of the disagreement between the pro-akratic and the anti-akratic. So let me explain how what I said in favor of AT* (as well as AT) interacts with what pro-akratic philosophers say in favor of the permissibility of epistemic akrasia. Recall, AT* says that:

> 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 *p* lacks *E* for one.

Hence, AT* claims that there couldn't be a particular kind of strong mismatch between a belief's positive epistemic status and its truth. The pro-akratic philosopher, however, thinks that there *could* be such a mismatch. Consider, for example, the following argument from Maria Lasonen-Aarnio for such a possibility:

> "Here is a very general line of argument. Let **F** be whatever property a doxastic state must have in order to be epistemically justified or rational. For at least numerous candidate properties **F**, a doxastic state D can have **F** even if another doxastic state—namely, the state of believing, or having high confidence, that D *lacks* **F**—itself has **F**.

[27] Another objection to Conjuncts come from the case Unbelievable Consequences, taken from Barnett (2021, p. 240): "Sylvie knows that the Oracle's past predictions all turned out true, so she rationally believes that today's prediction will be true. The Oracle then predicts: "You, Sylvie, will not just now come to believe any new conjunctions." This new prediction seems plausible enough, so hearing it does not affect the rationality of Sylvie's belief that the prediction is true." One may think that in such a case Sylvie should believe the conjuncts *today's prediction is that Sylvie won't now believe any new conjunctions*, as well as *today's prediction is true*, and yet Sylvie shouldn't believe their conjunction. Assuming this verdict is correct, it still doesn't seem to put too much pressure on Conjuncts. The reason is that this case is based on a peculiar feature of the supported proposition, namely it itself being about a conjunction, and this feature seems necessary for this kind of case to arise. This issue, however, seems orthogonal to the truth of Conjuncts, for if epistemic akrasia is permissible, then it should be possible for epistemic akrasia to be permissible for a subject satisfying Conjuncts in cases that don't satisfy the aforementioned feature. I thank a referee for another journal for bringing this objection to my attention.

> In fact, it is not at all easy to come up with a plausible candidate for **F** that rules out such a possibility. If, in general, false beliefs (or at least states involving high degrees of confidence in falsehoods) can have **F**, why not false beliefs about what doxastic states have that property?" (Lasonen-Aarnio 2020, p. 612)

Albeit being just one argument in the literature for the possibility of such a mismatch, I think it is illuminating to think about where this argument goes wrong, and what I have to say about it generalizes to other lines in the literature.[28] The argument fails for a reason that by now should be familiar: although a belief's positive epistemic status and its truth can come apart, there's a limit on such a mismatch. That there is *some* limit should be accepted by everyone. Recall AF, which states that for every positive epistemic status $E$ and for every proposition $p$, if one knows that $p$ is false, then $p$ lacks $E$ for one. AF poses a very clear limit on how far apart a belief with a positive epistemic status can come from that belief being true. Although a false belief can have any number of positive epistemic statuses, a belief known to be false cannot.

This further suggests that the connection between positive epistemic statuses and truth is more complicated than what comes out of Lasonen-Aarnio's argument. I think it is true that many simple proposals as to what grounds positive epistemic statuses won't exclude a radical mismatch between these very statuses and truth. However, AF gives us reason to believe that no such simple proposal can work anyway. Consider, for example, reliabilism about justification, and consider a belief that was in fact formed by a reliable process, but is false. This belief may be justified, but even the reliabilist should admit that it cannot remain justified once the subject comes to know that it is false. This knowledge of course doesn't affect the status of the original process as reliable.[29] After all, reliable processes need not be perfect processes. And yet, it affects the epistemic status of the belief.[30]

Given this, the reasons we have to admit some mismatch between a positive epistemic status $E$ and truth don't, on their own, generalize in the way suggested by Lasonen-Aarnio.[31] The

---

[28] For example, to Williamson (2011, 2014) and Horowitz (2014, 2019).

[29] Lasonen-Aarnio (2010) makes a convincing case for that.

[30] And indeed, due to issues having to do with higher-order evidence, Goldman (1979) amends his reliabilism as to allow for other facts, other than the actual reliability of the process by which a belief has been formed, to affect the justification of the relevant belief.

[31] Lasonen-Aarnio, as well as Williamson (2011, 2014) and Horowitz (2014, 2019), have taken different kinds of uncertainty to further support the possibility of a stronger mismatch between positive epistemic statuses and truth. However, one can acknowledge that the subject doesn't know what their evidence supports, for example, and still think that what their evidence in fact supports partly determines what they should believe about what their evidence supports. This could happen in at least two ways. First, it could happen because what attitude their evidence in fact supports with respect to some proposition $p$ constrains what attitude their evidence supports with respect to the proposition *my evidence supports* $p$, which in turn constrains whether they should believe that their evidence supports $p$. Second, it could happen because whether they should believe that their evidence supports $p$ is directly constrained by their evidence about $p$. This is in line with the thought that the subject may be required to believe that $p$ even if they don't know that their evidence supports $p$. (This is what happens in anti-luminosity cases.) So uncertainty doesn't directly support pro-akrasia; it does so only if further assumptions are made about the (lack of) connection between one's evidence and one's higher-order beliefs. And this is exactly what the preceding discussion in §1 is meant to address.

above discussion shows that although we can justifiably believe in some falsehoods, there are still limits on how far apart justification and truth can come apart. And the same is true for other positive epistemic statuses. Thus, the reasons for accepting the first kind of mismatch don't support the stronger mismatch that is needed for the permissibility of epistemic akrasia.[32]

Here's another, related way to think of the dialectic. Lasonen-Aarnio argues that cases of epistemic akrasia teach us that we can't hold on to both evidentialism and the impermissibility of epistemic akrasia, and that thus, we should keep the former and give up the latter. We can think of the discussion supporting AT* and AT as presenting a way of thinking of positive epistemic statuses, such as evidential support, that doesn't force us to decide between evidentialism and the impermissibility of epistemic akrasia. Although some purported cases of permissible epistemic akrasia presented in the literature have a considerable appeal, a closer look at the nature of positive epistemic statuses teaches us that these cases are impossible: it is never the case that your evidence supports both *p* and that your evidence does not support *p*. If evidential support could have been structured in this way, evidential support wouldn't have been a good guide to the truth. And if cases like this cannot arise in principle, then we can hold on both to principles like AT* and AT (and consequently, the impermissibility of epistemic akrasia) and to evidentialism. Moreover, the discussion shows that this response isn't ad hoc; rather, it stems from a very natural picture of the nature and role of positive epistemic statuses—a picture on which there's a strong connection between evidential support and truth.

Now, of course my opponent can insist on an alternative way of understanding positive epistemic statuses, one on which AT* and AT do come at the expense of evidentialism. But that would be akin to a reliabilist who insists that in cases where the subject first forms the belief that *p* using a reliable method, and then comes to know that *p* is false, the subject is still justified in believing that *p*. This amounts to rejecting AF. A more plausible approach to cases like this would be to keep AF, and instead revise one's views of what justification consists in—maybe by accepting a more nuanced form of reliabilism, as people like Goldman in fact do. I admit that AF is even more intuitive than AT* and AT, and so rejecting the latter is not as unreasonable as rejecting the former. But I think that discussing AF helps us see that we must have nuanced views about what evidential support and justification consist in—ones that take very seriously the connection between these statuses and truth. And once we acknowledge this, it's hard to see why we should prefer accepting the permissibility of epistemic akrasia over accepting AT* and AT.

---

[32] This is the same line I take in response to cases such as Williamson's Unmarked Clock (Williamson 2011). Of course, proponents of epistemic akrasia can come up with models that invalidate principles such as AT and AT*. In particular, the Unmarked Clock can be thought of as a model that invalidates AT* as applied to $E$ = evidential support. However, this section gives reasons to think that the assumptions of these models are unmotivated, and hence that we shouldn't think these models can be true of positive epistemic statuses.

## §4 Conclusion

I've just argued that one ought not be epistemically akratic on the basis of epistemic considerations alone. The argument showed that taking epistemic akrasia to be permissible is not only counterintuitive, but also conflicts with there being a suitably strong connection between positive epistemic statuses and *truth*. I want to conclude the paper by drawing an analogy between Moore's paradox and epistemic akrasia. The analogy, I believe, sheds light on my treatment of epistemic akrasia, and lends it some more support.[33]

Consider the proposition at the heart of Moore's paradox: *p, but I don't believe p*. Both the Moorean proposition and the akratic conjunction (AC) exhibit the following features: (1) both are consistent propositions—both conjuncts of each conjunction can be true at the same time; (2) both intuitively sound like something no subject should believe when applied to their current self; and (3) both are such that a subject can permissibly believe them when applied to other subjects, or to themselves at other times. It seems to me that the same kind of issue is at play when we consider what the right doxastic attitude towards the Moorean proposition (*p, but I don't believe p*) and towards AC is.

Consider the following case: a highly reliable oracle tells me, for some proposition *p* that I haven't considered before, that *p*, but that I don't believe *p*, nor will I believe *p* in the next five minutes. Since the oracle is highly reliable, it seems like I may believe *p, but I don't believe p*, or have high credence in this proposition (maybe one that matches the oracle's reliability). However, it's quite clear that I ought not believe this conjunctive proposition. This is because believing it is self-undermining: I can easily come to know that if I believe it, it is false. (Notice that if this is true, it is also true that I can come to know that if my belief has $E$, it is false, and so we can use AKT to infer the impermissibility of believing the Moorean proposition.)

Or, in the akratic case, consider: a highly reliable oracle tells me that *q* and that my (potential) belief in *q* lacks $E$ (now or for the next five minutes). Since the oracle is highly reliable, it seems like I may believe *q, but q lacks E for me*, or have high credence in that proposition (maybe one that matches the oracle's reliability). But, in light of what I have argued, it is clear that I ought not believe this conjunctive proposition. This is because believing it is self-undermining: I can easily come to know that if my belief in it has $E$, it is false.

The cases of epistemic akrasia and Moore's paradox are therefore analogous. In both cases, it seems like one may believe each of the conjuncts of the relevant conjunctions. After all, oracles give one excellent evidence. But in both cases, considering both conjuncts together in the form of conjunction reveals that believing those conjunctions conflicts with there being

[33] Smithies (2012) also takes the cases of Moore's paradox and epistemic akrasia in its justification version (one believes *q* and that *one isn't justified in believing that q*) to be closely related, although his treatment of the two relies on the claim that justification is accessible. See also Sorensen (2002) for the similarity between Moore's paradox and other self-effacing paradoxes.

a suitably strong connection between positive epistemic statuses and truth. Hence, both cases teach us something important about the nature and structure of positive epistemic statuses.[34]

## Acknowledgements

This paper has benefited from the help of many people. Special thanks to Jane Friedman, Richard Roth and Patrick Wu for many cycles of comments and discussions. I also thank Cristina Ballarini, Evan Behrle, David Chalmers, Tez Clark, Cian Dorr, Kevin Dorst, Jorge Ferreira, Noga Gratvol, Ayala Colette Haddad, David Hills, Adam Lovett, Matthew Mandelkern, Jessica Moss, William Nava, Antonia Peacocke, Jim Pryor, Ali Rezaei, Miriam Schoenfield, Roger White, Crispin Wright and Snow Zhang.

## References

Barnett, David James (2021). "Self-Knowledge Requirements and Moore's Paradox." *Philosophical Review* 130(2): 227–262.

Christensen, David (2016). "Disagreement, Drugs, etc.: from Accuracy to Akrasia." *Episteme* 13(4): 397–422.

Coates, Allen (2012). "Rational Epistemic Akrasia." *American Philosophical Quarterly* 49(2): 113–124.

Conee, Earl (1987). "Evident, but Rationally Unacceptable." *Australasian Journal of Philosophy* 65(3): 316–326.

Dorst, Kevin (2020). "Evidence: A Guide for the Uncertain." *Philosophy and Phenomenological Research* 100(3): 586–632.

Fantl, Jeremy and Matthew McGrath (2002). "Evidence, Pragmatics, and Justification." *Philosophical Review* 111(1): 67–94.

Foley, Richard (1979). "Justified Inconsistent Beliefs." *American Philosophical Quarterly* 16(4): 247–257.

Goldberg, Sanford C. (2017). "Should Have Known." *Synthese* 194: 2863–2894.

Goldman, Alvin I. (1979). "What is Justified Belief?" In George S. Pappas (ed.), *Justification and Knowledge* (D. Reidel Publishing Company), pp. 1–23.

Greco, Daniel (2014). "A Puzzle about Epistemic Akrasia." *Philosophical Studies* 167(2): 201–219.

Horowitz, Sophie (2014). "Epistemic Akrasia." *Nous* 48(4): 718–744.

— (2019). "Predictably Misleading Evidence." In Mattias Skipper & Asbjørn Steglich-Petersen (eds.), *Higher-Order Evidence: New Essays* (Oxford University Press), pp. 105–123.

Joyce, James M. (1998). "A Non-Pragmatic Vindication of Probabilism." *Philosophy of Science* 65(4): 575–603.

Lasonen-Aarnio, Maria (2010). "Unreasonable Knowledge." *Philosophical Perspectives* 24:1–21.

— (2020). "Enkrasia or Evidentialism? Learning to Love Mismatch." *Philosophical Studies* 177(3): 597–632.

Le Morvan, Pierre (2012). "On Ignorance: A Vindication of the Standard View." *Philosophia* 40: 379–393.

Littlejohn, Clayton (2012). *Justification and the Truth-Connection*. Cambridge University Press.

— (2018). "Stop Making Sense? On a Puzzle about Rationality." *Philosophy and Phenomenological Research* 96(2): 257–272.

Nelkin, Dana K. (2000). "The Lottery Paradox, Knowledge, and Rationality." *Philosophical Review* 109(3): 373–409.

Neta, Ram (2019). "The Puzzles of Easy Knowledge and of Higher-Order Evidence: A Unified Solution." In Mattias Skipper & Asbjørn Steglich-Petersen (eds.), *Higher-Order Evidence: New Essays*, Oxford University Press, pp. 173–188.

Ryan, Sharon (1991). "The Preface Paradox." *Philosophical Studies* 64(3): 293–307.

Schoenfield, Miriam (2015). "Bridging Rationality and Accuracy." *Journal of Philosophy* 112(12): 633–657.

Smithies, Declan (2012). "Moore's Paradox and the Accessibility of Justification." *Philosophy and Phenomenological Research* 85(2): 273–300.

[34] See Conee (1987) and Neta (2019) for related points about the structure of acceptance and the structure of evidential support respectively.

Sorensen, Roy A. (2002) [1988]. *Blindspots*. Oxford University Press.

Titelbaum, Michael G. (2015). "Rationality's Fixed Point (Or: In Defense of Right Reason)." *Oxford Studies in Epistemology* 5: 253–294.

Weatherson, Brian (2019). *Normative Externalism*. Oxford University Press.

Williamson, Timothy (2000). *Knowledge and Its Limits*. Oxford University Press.

— (2011). "Improbable Knowing." In Trent Dougherty (ed.), *Evidentialism and Its Discontents* (Oxford University Press), pp. 147–164.

— (2014). "Very Improbable Knowing." *Erkenntis* 79: 971–999.