Accuracy and epistemic conservatism

Florian Steinberger
Birkbeck College, University of London

December 15, 2018

Abstract: Epistemic utility theory (EUT) is generally coupled with *veritism*. Veritism is the view that truth is the sole fundamental epistemic value. Veritism, when paired with EUT, entails a methodological commitment: Norms of epistemic rationality are justified only if they can be derived from considerations of accuracy alone. According to EUT, then, believing truly has epistemic value, while believing falsely has epistemic disvalue. This raises the question as to how the rational believer should balance the prospect of true belief against the risk of error. A strong intuitive case can be made for a kind of epistemic *conservatism*—that we should disvalue error more than we value true belief. I argue that none of the ways in which advocates of veritist EUT have sought to motivate conservatism can be squared with their methodological commitments. Short of any such justification, they must therefore either abandon their most central methodological principle or else adopt a permissive line with respect to epistemic risk.

1 Introduction

Advocates of epistemic utility theory (EUT) generally espouse the following two claims. First, belief aims at accuracy. Second, norms governing belief are subordinate to the aim of achieving accuracy. The first claim is generally understood as a commitment to monism about fundamental epistemic value and to truth as the sole epistemic value.

Let us employ the standard label *veritism* for this position. The second claim expresses a methodological commitment: Norms of epistemic rationality are justified only to the extent to which they are conducive to accuracy. 'We should not impose any external constraints on the class of doxastic alternatives' other than the ones that 'can be derived purely from considerations of accuracy' (Dorst Forthcoming: 6). The second claim follows from the first in conjunction with a broadly consequentialist conception of norm justification inherent in EUT.[1]

Against this value-theoretic framework, the rational believer aims to have a doxastic state that best promotes epistemic value. To arrive at such a state she employs the tools of decision theory. Just as in practical decision theory this requires that we define a utility function that specifies outcomes for each of one's possible actions depending on the state of the world. In the context of EUT, the 'choice' in question is one between different doxastic states. An epistemic utility function tracks our performance by assigning positive epistemic utility to true belief and negative epistemic utility to false belief. But here we face an important question: How does the value of true belief relate to the disvalue of false belief? Do we stand to gain more by believing truly than we stand to lose by believing falsely? Is it the other way around? Or is value of the former equal the value of the latter? Let R represent the value of "getting it right" and $-W$ represent the disvalue of "getting it wrong". There are then three possible stances towards epistemic risk:

Radicalism: $R > W$

Centrism: $R = W$

[1]We do well to distinguish clearly EUT from veritism. Although the vast majority of the discussions of EUT concern the conjunction of the two views, neither one entails the other. One might apply the EUT framework to promote a different epistemic value (e.g. knowledge). Conversely, one might be a veritist, while, for instance, rejecting the epistemic consequentialism inherent in EUT (see e.g. Sylvan 2018). That said, my target here is the combination of EUT and veritism. For simplicity, I henceforth use 'veritist EUT' or just 'EUT' to designate it. I am grateful to an anonymous referee for urging me to clarify this point.

Conservatism: $W > R$

The choice is a familiar one. William James famously maintained that in pursuing accuracy we are subject to the double imperative: 'Believe truth! Shun error!' Working out what to believe involves weighing the risk of error against the benefit of believing truly. At the limit, a believer who cared only about maximizing true beliefs would indiscriminately believe any proposition whatsoever. At the other extreme, a believer who above all else sought to avoid error would refrain from believing anything at all. Neither option is tenable; both R and W must receive *some* weight. The challenge is to find the optimal balance between the two.

James himself believed that one's attitude towards epistemic risk is largely a matter of one's 'passional nature', one's intellectual temperament. It is not clear, though, that such a liberal attitude is warranted. There are compelling reasons to espouse conservatism. For suppose we are deliberating over whether to believe p or its evidentially equally supported negation $\neg p$. Intuitively, we have no grounds for believing either of the two: We should suspend. Our intuitions align with conservatism. But if $R > W$, we have reason to be epistemic risk takers and so to go out on a limb and to believe one and disbelieve the other, even in the absence of evidence. Centrists, equally implausibly, are indifferent between believing one of the two, believing (or disbelieving) both and suspending.[2]

If such a strong case can be made for risk aversion, the proponent of veritist EUT should want to be on board. And indeed she is. Advocates of EUT argue for conservativism. But—and this is the crux of the paper—there would seem to be no purely accuracy-based case for conservatism within the EUT framework. In the absence of any such argument, proponents of EUT must therefore either abandon their methodological principles or else bite the bullet and embrace a Jamesian pluralism about epistemic risk.

The plan is this. In §2, I provide a brief overview of EUT. §3 looks at two intuitively

[2] I discuss this case in greater detail in §3.

compelling arguments for epistemic conservatism and finds them wanting by EUT's own methodological standards. A more sophisticated argument due to Kenneth Easwaran (2016) is considered but ultimately also rejected in §4.

2 EUT—a sketch

Our challenge, we said, is to strike the optimal balance between maximizing true beliefs and minimizing false ones. EUT proposes to bring the tools of decision theory to bear on the problem. In this section I provide a brief overview of the standard framework for full belief.[3]

We suppose our agent is entertaining a finite set of propositions $\mathcal{P}$. We may assume that whenever $A \in \mathcal{P}$, $\neg A \in \mathcal{P}$, but $\mathcal{P}$ need not be closed under Boolean operations. For each proposition in $\mathcal{P}$, the agent may adopt one of three attitudes: belief (B), disbelief (D) or suspension of belief (S). Formally, we can represent the agent's various possible 'choices' over $\mathcal{P}$ as *belief functions* of the form:

$$b : \mathcal{P} \to \{B, S, D\}$$

We assess belief functions—no surprises here—based on their accuracy. Our standard of accuracy is the relevant proposition's truth-value at a given possible world, where the set of possible worlds, $W_{\mathcal{P}}$, is generated by $\mathcal{P}$. A possible world is simply represented by means of a consistent valuation function

$$w : \mathcal{P} \to \{t, f\}.$$

Each choice of a doxastic attitude is then thought to have a numerically represented epistemic utility relative to a world. This is our epistemic utility function, which, for

[3]Though my presentation may differ from others' presentations on points of detail, these differences are immaterial for present purposes.

any attitude-truth-value pair, returns the appropriate epistemic utility:

$$eu : \{B, D, S\} \times \{t, f\} \to [-\infty, \infty]$$

That is, for any proposition $A \in \mathcal{P}$, eu returns a score as a function of one's attitude towards it and the proposition's truth-value at the world in question. This is where R and W enter the picture: R represents the score for holding an accurate attitude; $-W$ the penalty one incurs for holding an inaccurate attitude. Suspending is assumed to yield a neutral score.[4] Hence:

$$eu(B, t) = eu(D, f) = R$$

$$eu(S, t) = eu(S, f) = 0$$

$$eu(B, f) = eu(D, t) = -W$$

Epistemic utility is generally assumed to be additive. (This assumption is not essential. See e.g. Dorst Forthcoming.) The overall epistemic utility, EU, of a belief function b at a world w is then determined in the obvious way

$$EU(b) = \sum_{A \in \mathcal{P}} eu(b(A), w(A)).$$

Of course our aim is to match our beliefs to the actual world, not just any old world. Hence, we seek to maximize the *actual* epistemic utility of our beliefs. Alas, we are not in general in a position to know which of the doxastic possibilities is actual. Nevertheless, even in our state of ignorance, there are certain belief functions that we can immediately eliminate from consideration. A belief function that has less epistemic utility than another *however* the world turns out to be, plainly, is not a candidate. It

[4] To maintain the distinction between the attitude of suspending belief and a lack of any doxastic attitude altogether, we may assume that $\mathcal{P}$ is a proper subset of a larger set of propositions. See Friedman 2013.

would be irrational for any agent to adopt it. We can capture this type of irrationality in terms of the familiar notion of dominance.[5]

> **Strict Dominance:** A belief function b is strictly dominated by a belief function b' iff, for all worlds w, $EU(b', w) > EU(b, w)$.

> **Weak Dominance:** A belief function b is weakly dominated by a belief function b' iff, for all worlds w, $EU(b', w) \geq EU(b, w)$, and there exists a world w' such that $EU(b', w') > EU(b, w')$.

Notice that a belief function b's being dominated by another (un-dominated) function b' only shows that it would be irrational to opt for b. It says nothing about which belief function one should adopt (in particular, it does not in general recommend adopting b').

The notion of dominance so applied naturally gives rise to a potential norm of epistemic rationality. Following Easwaran (2016), we call a belief function b *strongly coherent* just in case it is not even weakly dominated, and that b is *weakly coherent* just in case it is not strongly dominated. The following rationality requirement thus falls right out of EUT's central commitments:

> **Strong Coherence:** One ought to have strongly coherent beliefs.

Dominance reasoning is available to us even in the absence of any information regarding the likelihood of the various doxastic possibilities. Imagine now that we do acquire such information. More precisely, assume there is a probability function P which assigns a probability weighting to each of our possible worlds. Given such a measure of likelihood, we can calculate the *expected* epistemic utility of each of the belief functions available to us:

$$EEU_P(b) = \sum_{w \in W} P(w) EU(b, w)$$

[5] One option's being dominated by another makes choosing it irrational only if the dominating option is not itself dominated. See Pettigrew 2016a. I am setting aside a number of further complications here. In particular, for simplicity, I am assuming act-state independence.

Note that our definition is neutral with respect to the interpretation of our probability function.

Maximizing expected epistemic utility (EEU) is of course a decision principle in its own right. And it may be thought to give rise to a separate requirement of epistemic rationality. Moreover, so long as we maximize EEU relative to a regular probability function—one that assigns every world a non-zero probability—doing so is a sufficient condition for strong coherence. What is more, if $W \geq R$, it can be shown that a belief function maximizes expected utility just in case there exists a probability function such that for any proposition A (see (Easwaran, 2016: 828) for details.):

- $b(A) = B$ iff $1 \geq P(A) \geq \frac{W}{R+W}$
- $b(A) = S$ iff $\frac{W}{R+W} \geq P(A) \geq \frac{R}{R+W}$
- $b(A) = D$ iff $\frac{R}{R+W} \geq P(A) \geq 0$

If $W < R$, a belief function maximizes expected utility just in case there exists a probability function such that for any proposition A:

- $b(A) = B$ iff $1 \geq P(A) \geq \frac{1}{2}$
- Suspending never has maximal expected utility.
- $b(A) = D$ iff $\frac{1}{2} \geq P(A) \geq 0$

It is noteworthy that if P is interpreted as the agent's credence function, this result yields a version of the so-called *Lockean Thesis* according to which fully (rationally) believing is (in a sense to be made precise) having credence in excess of a threshold. The threshold is set by the appropriate ratio between R and W.[6]

[6]Dorst (2017) argues for a more sophisticated variable-threshold Lockean account, which he takes to provide a metaphysical reduction of full belief to credence. Easwaran (2016), by contrast, argues in favor of the primacy of full belief.

So much for the overview of the framework. It is not hard to see that the question of the relative values of R and W is central to the determination of precisely what EUT maintains the norms of epistemic rationality consist in. To see how we now take a closer look at the case for conservatism.

3 The coin flip case

Consider the following case (see Dorst Forthcoming). You flip a fair coin. Let p be the proposition that the coin lands heads. What attitude is it rational for me to adopt with respect to p in absence of any further evidence? The answer, we said, seems obvious: I should suspend. But consider the following decision matrix. The two left most columns represent the propositions' truth values at the two possible worlds. The remaining columns represent the score associated with the various belief functions. For instance 'BB' represents the belief function that believes both p and its negation. The cells associated with each belief function represent the epistemic utility of the belief function at a world. For example, the first cell in the BB column corresponds to $EU(b, w) = eu(b(p), w(p)) + eu(b(\neg p), w(\neg p))$ where b is the belief function that believes both p and $\neg p$ and w is the world at which p is true and its negation false.

p	$\neg p$	BB	BD	SS	BS
t	f	R-W	2R	0	R+0
f	t	-W+R	-2W	0	-W+0

(The remaining cases—DD, DB, SB, DS, SD—are strictly analogous and so can be omitted here and in the following.) Clearly, our assessment of the various options depends on our attitude towards epistemic risk. If, as a radical, my aim is coherence, suspending is not rationally permissible. Given that $R > W$, believing both p and $\neg p$ dominates believing neither. BB and BD can both be shown to be coherent because they maximize

EEU. If $R = W$, then BB, BD and SS all maximize EEU and so are all coherent. Finally, if $W > R$, suspending is the only coherent option.

Conservatism clearly seems—*pace* James—the sensible position. Dorst concurs:[7]

> does she [a rational agent] believe [the fair coin will] land heads? Or tails? Or both? Or neither? Clearly neither. But if she cared more about seeking truth than avoiding error, why not believe both? She'd then be guaranteed to get one truth and one falsehood, and so be more accurate than if she believed neither. Yet believing both is *not* as accurate as believing neither—so belief is conservative (Dorst Forthcoming).

But what makes suspending more accurate? Dorst does not tell us. Indeed, it is not clear what accuracy-based reasons there might be for thinking so. After all, accuracy is measured via the epistemic utility of each of the options, but as Dorst himself points out 'we only get a metric once we weigh these factors [that of seeking truth and that of avoiding falsity] against each other (10)'. In other words, we must first settle on the relative values of R and W *before* we are able to keep score and so are able to pronounce on matters of accuracy. I suspect that this is ultimate reason why veritist EUT is unable to make a case for conservatism.

It is tempting to brush such doubts aside as sophistry, so powerful are our intuitions in favor of conservatism. But what exactly are the considerations that undergird our intuitions, and why should the proponent of EUT not be in a position to appeal to them? We can identify at least two such considerations: One concerns the impermissibility of believing contradictions; the other the impermissibility of believing on the basis of insufficient evidence.

[7]One finds very similar arguments in other recent significant work. Easwaran (2012: 824) deems the upshot the consequences of adopting epistemic radicalism 'strange enough' to warrant the blanket stipulation $W \geq R$. However, Easwaran goes on to argue that this stipulation is not 'substantive'. I discuss this further argument in §4. Shear and Fitelson (forthcoming) similarly endorse conservativism on account of the counterintuitive consequences of the alternatives.

According to the first line of thought, we should not waste our time with any metric that licenses belief in a proposition and its negation. Believing both A and $\neg A$ is epistemically impermissible for the obvious reason that at least one of my beliefs is bound to be false.[8] Thus, in consciously believing a contradiction, I knowingly hold beliefs that are less accurate than they might be. The reasoning is reminiscent of the dominance reasoning that grounds the notion of strong coherence. There is a subtle difference in the form of a quantifier shift, though. Dominance reasoning has the form:

> (A) If there exists a belief function b' such that for every possible world w, b' outperforms b at w, then it is irrational to adopt b.[9]

The present case for non-contradictory beliefs has the form:

> (B) If, for every possible world w, there exists a belief function b' such that b' outperforms b at w, then it is irrational to adopt b.

But the advocate of EUT rejects (B). For if (B) were sufficient for irrationality, then inconsistency too would count as irrational. Yet, advocates of EUT are clear that rational belief may be inconsistent (both strong coherence and maximizing EEU are compatible with inconsistency) as in the case of lottery or the preface paradoxes.

Of course many authors maintain that there is a relevant difference between the principles of non-contradiction and of consistency. (See e.g. Foley 1979.) Given the Lockean thesis, it might be argued, full belief follows (so to speak) rational credence. Hence, while one may rationally have sufficiently high credence in each member of an inconsistent set of propositions, one cannot rationally have high credence in each of a pair of contradictory propositions. Therein, according to this attempted response, lies the disanalogy. And this much, surely, is true. But this is of no help to veritist EUT. As we noted above, it suffices that my credence in both A and $\neg A$ is 0.5. So long as it

[8] I am setting dialetheism aside here, which would require various qualifications.

[9] Assuming again that b' is not itself dominated by a further option.

is assumed that $R > W$, believing both A and its negation is still a way of maximizing expected epistemic utility and so is in line with the Lockean thesis.

Turn now to the second, evidentialist case for conservatism. Again the thought is as simple as it is plausible (at least at first blush). We should believe neither that the fair coin lands heads nor that it lands tails because we have no evidence either way. This is just Hume's dictum exhorting us to 'proportion our beliefs to the evidence'. While a broadly evidentialist norm may be thought to have some initial appeal, the question from the point of view of veritist EUT is how such a norm may be justified with reference to the aim of accuracy alone. Following Pettigrew (2013), we can think of evidential norms as relating to accuracy in a variety of ways. An evidential norm might be thought to act as a side-constraint on our accuracy-oriented choice of beliefs. Alternatively, we might think of evidential support as a separate value alongside that of accuracy, which may or may not be commensurable with it.[10] By contrast, Pettigrew remains true to the core commitments of veritist EUT and attempts to show how evidential norms fall out of the aim of accuracy. It is doubtful that it really is possible to derive evidential norms purely on the basis of accuracy.[11] But even if we grant that it is possible, as Pettigrew (2013) argues, to derive evidential norms governing credences such as (a version of) the Principal Principle, the Principle of Indifference and Conditionalization on the basis of purely alethic factors together with various decision principles in a methodologically sound manner, this still does not vindicate our intuitions regarding the coin flip case: Since the coin is fair, the chance of its landing heads is 0.5. Given the Principal Principle my credence in p should also be 0.5. Seeing that my credences are assumed to be probabilistically coherent, my credence in $\neg p$ too is 0.5. And yet, as a radical I can continue to believe both p and $\neg p$ while maximizing EEU.[12] The Principle of Indifference

[10] Following the latter approach, Easwaran and Fitelson (2015) offer an illuminating discussion of evidential norms and their role in motivating coherence norms in an EUT-style framework. For all its merits, though, their approach espouses a kind of value pluralism and so abandons veritism.

[11] See (Meacham 2016) for a lucid discussion of the obstacles faced by such approaches.

[12] It is easy to check that if $R > W$, a belief function that believes both p and $\neg p$ is not dominated

requires that I regard p and $\neg p$ as equiprobable. It does not fix the relative values of R and W.

In summary, the proponent of EUT is unable to appropriate either of these plausible strategies for vindicating our intuition in the coin flip case.

4 Risk aversion as a modelling assumption

Easwaran (2016) offers a more sophisticated argument with the aim of showing that setting $W > R$ is merely a 'naming convention' and not a 'substantive assumption'. It can be summarized as follows. Suppose we have two attitudes B and S. Let b_T be the value of having attitude B towards a true proposition, and b_F the value of having B towards a false proposition. Likewise for S. Easwaran reasonably assumes that neither attitude should dominate the other. That is, neither attitude should receive a higher score than its alternative regardless of the truth-value of the proposition to which they are borne. In short, we may exclude the possibility that $b_T \leq s_T$ and $b_F \leq s_F$ or vice versa. It follows by simple logic that $b_T > s_T$ and $b_F < s_F$ or vice versa. Since we have not made any assumptions about our two attitudes, we may stipulate that $b_T > s_T$ and $b_F < s_F$. Thus, B is the attitude that yields a higher score when applied to a true proposition and S the attitude that fares better when applied to a false proposition. A natural interpretation given my choice of labels is to think of B as belief and S as suspension. However, for all we have said, we could just as easily interpret B as suspension and S as disbelief. To avoid confusion, let us introduce D for disbelief and S' for the alternative notion of suspension associated with disbelief. We then have $s'_T > d_T$ and $s'_F < d_F$. Assuming that one believes a proposition just in case one disbelieves its negation, we can see the two pairs of attitudes to be related as follows (where 'a' refers to an agent):

according to the other decision principles (Maximin, Chance Dominance) either.

- a bears B to $\neg A$ iff a bears D to A.

- a bears S to $\neg A$ iff a bears S' to A.

It is natural to set $b_T = d_F = R$ and $b_F = d_T = -W$, and $s_T = s'_F = R'$ and $s_F = s'_T = -W'$. We know that $R > R'$ and $W' > W$. As Easwaran points out, the scores are invariant under linear transformations. None of the inequalities are affected by adding or multiplying both sides by the same values (modulo the obvious adjustments necessary when multiplying both sides by a negative value). By rescaling, we can ensure that $R' = W' = 0$.[13]

How, then, can we distinguish between the two pairs of attitudes? By fiat, assuming, along with Easwaran, that attitudes are individuated solely by the way they contribute to doxastic states (2016: 838). The score of bearing B (S) to both A and $\neg A$ is $b_T + b_F$ ($s_T + s_F$). So the question which of B and S is better applied to a pair of contradictory propositions is the question whether $b_T + b_F \geq s_T + s_F$ or $b_T + b_F < s_T + s_F$, and again with some rescaling: $b_T + b_F \geq 0$ or $b_T + b_F < 0$? This, in turn, is tantamount to asking whether $R \geq W$ or $R < W$. The analogous question can be asked about S' and D. Easwaran now stipulates that $b_T + b_F \leq s_T + s_F$ (i.e. that $R \leq W$), but that $s'_T + s'_F \geq d_T + d_F$ (i.e. that $R \geq W$). The relative value between R and W, then, is nothing more than a naming convention to distinguish the otherwise equivalent descriptions (in terms of B, S or in terms of S', D) of a doxastic state.

How does this work in practice? Well,

> imagine that we start with one attitude that gets score R if the proposition it is applied to is true, and score $-W$ if the proposition is false, and another attitude that gets score 0 in either case. If $R > W$ [...], then our conventions from the previous section mean that we should label the first attitude as A' and the second attitude as B' (2016: 838).

[13]Note that we could equally have set R and W equal to zero, provided we make the appropriate compensating adjustments.

A' and B' correspond to our S' and D respectively. Thus, following Easwaran's convention, the would-be belief should be re-described as S' (suspension or lack of disbelief) and the attendant attitude receiving the score 0 as D (disbelief). To do so we must rescale. We have $d_F = b_T = R$ and $d_T = b_F = -W$. Hence, to ensure that $d_F = d_T = 0$ we must add $-R$ to d_F and W to d_T. Adding the same amount to the other side of the inequalities yields: $s'_T = W$ and $s'_F = -R$. If we set $W = R'$ and $-R = W'$, we find that the attitude that is now playing the role of belief does indeed satisfy the requirement that $W' > R'$.

The problem with this argument is twofold. First, it makes the *substantive* assumption that $R \neq W$ (idem, p. 842). Second, it relies on problematic prior assumptions. Easwaran's framework operates only with two attitudes: belief and lack of belief. In a further appendix (C), Easwaran seeks to demonstrate that frameworks that incorporate additional states (viz. suspension and disbelief) come to the same thing. His argument proceeds by first assuming we operate with three attitudes X, Y and Z (intuitively: belief, suspension and disbelief). We again assume that no attitude is dominated by the others. We may assume, furthermore, that B is the 'positive', S is the 'neutral' and D is 'negative' attitude (i.e. $x_T > y_T > z_T$ and $x_F < y_F < z_F$). Easwaran then argues that these three attitudes collapse into two attitudes (B and S) provided we set $b_T = x_T + z_F$, $b_F = x_F + z_T$ and $s_T = s_F = 0$ (the latter via linear transformations if necessary). Easwaran's discussion is illuminating. It does, however, rely on the *symmetry condition.* The symmetry condition posits that an agent can have X towards A if and only if she has Z towards $\neg A$. This seems innocent, but it means that for any pair of propositions of the form $(A, \neg A)$, the agent can only take one of three pairs of attitudes $(X, Z), (Y, Y)$ or (Z, X). The possibility of believing both A and $\neg A$ is therefore simply stipulated out of existence. Easwaran comments:

> it seems plausible that there is a further requirement that one ought to have a doxastic state that satisfies the symmetry condition. Thus, for three-attitude

> doxastic states, we can define coherence in terms of dominance together with the symmetry requirement. Thus, although there will be three-attitude doxastic states that don't correspond to any two-attitude doxastic states, all of the coherent ones will correspond (2016: 843).

Easwaran's proposal to build the symmetry requirement into the notion of coherence is not implausible *per se*. It is not, however, true to EUT's methodological commitments. This formulation of the coherence norm is no longer motivated solely by accuracy (and decision-theoretic) considerations alone, and the inbuilt norm ruling out contradictory beliefs once again lacks motivation. However, without the symmetry requirement the formal equivalence of the three-attitude and the two-attitude framework does not hold. Since the previous argument presupposed this equivalence, the assumption that $R < W$ *is* substantive (and hence unjustified).

A possible reaction to this might be relax the veritist EUT's methodological strictures. Perhaps the guiding principle endorsed by Dorst and others that only such epistemic norms are permissible as can be 'derived from accuracy alone' is too demanding.[14] For instance, one might try to motivate a prohibition against contradictory belief on the basis of a particular conception of the nature of belief, thereby vindicating conservatism via Easwaran's argument. Of course such a metaphysical account of belief would have to be independently defended against alternative accounts that make incompatible claims about the constitutive norms governing belief (e.g. accounts compatible with dialetheism or interpretivist accounts that would impose stronger logical closure and consistency constraints). More generally it might be thought that the advocate of EUT should be able to help herself to plausible epistemic norms even in the absence of a truth-oriented decision-theoretic justification. That is, it might be thought that the proponent of veritist EUT should espouse a weaker methodological policy according to which justifiability in veritist decision-theoretic terms would be a sufficient though not

[14] I am grateful to an anonymous referee for pressing me on this point.

a necessary condition for epistemic norm-hood—other norms may be admissible so long as they do not contravene EUT-backed norms. Advisable or not, it should be noted that such an innocent-until-proven-guilty approach to epistemic norms would constitute a significant departure from what Dorst describes as the 'Accuracy First methodology':

> we should leave open a wide class of doxastic alternatives, since we want to derive norms from accuracy alone. Again, Accuracy First is rarely stated, but it's implicit in the way epistemic utility theorists set up their frame-works and paint the big picture of what they are trying to do (Dorst Forthcoming: 6).

As the above example of an attempt to derive the norm of non-contradiction from an account of the nature of belief makes plain, such a methodological reorientation would put an end to EUT as a monolithic enterprise. There would be a multitude of different decision-theoretic veritist enterprises each with its own set of optional non-EUT-based epistemic norms.

5 Conclusion

The moral of this story, then, is that EUT has so far been unable to provide a purely accuracy-based justification of conservatism. Consequently, EUTers owe us such a justification. So long as they are unable to provide one, they must abandon the Accuracy First methodology, or else join James in adopting a permissive line towards epistemic risk.[15][16]

[15] See Pettigrew 2016b for an interesting discussion of Jamesian epistemology from the point of view of EUT.

[16] I would like to thank two anonymous referees and an associate editor for valuable comments and suggestions. Moreover, I am grateful to Julien Dutant and audiences in Bonn, Cologne, Hamburg and Birkbeck College for helpful discussions.

References

[1] K. Dorst. Forthcoming. Lockeans maximize expected accuracy. *Mind.*

[2] K. Easwaran. 2016. Dr. Truthlove, or, how I learned to stop worrying and love Bayesian probability. *Noûs* 50: 816–53.

[3] K. Easwaran and B. Fitelson. 2015. Accuracy, coherence, and evidence. In T. Szabo Gendler and J. Hawthorne, editors, *Oxford studies in epistemology* volume 5: 61–96.

[4] R. Foley. 1979. Justified inconsistent beliefs. *Americal philosophical quarterly* 16: 247–57.

[5] J. Friedman. 2013. Suspended judgment. *Philosophical studies* 162: 165–81.

[6] C. J. G. Meacham. Forthcoming. Can all-accuracy accounts justify evidential norms? In K. Ahlstrom-Vij and J. Dunn, editors, *Epistemic consequentialism.* Oxford: Oxford University Press.

[7] R. Pettigrew. 2013. Accuracy and evidence. *Dialectica* 67: 579–96.

[8] R. Pettigrew. 2016a. *Accuracy and the laws of credence.* Oxford: Oxford University Press.

[9] R. Pettigrew. 2016b. Jamesian epistemology formalised: An explication of 'The will to believe'. *Episteme* 3: 253–68, 2016b.

[10] R. Pettigrew. 2017. Epistemic utility and the normativity of logic. *Logos and episteme*: 455–92.

[11] T. Shear and B. Fitelson. Forthcoming. Two approaches to belief revision. *Erkenntnis.*

[12] K. Sylvan. 2018. Veritism unswamped. *Mind*, 127: 381–435.