

Kristoffer Ahlstrom-Vij and Jeffrey Dunn (eds.), *Epistemic Consequentialism*.
Oxford: Oxford University Press, 2018. vii + 335 pp.

Epistemic consequentialism is the view that (at least some) features of epistemic rationality can be explained in terms of features of epistemic value. The standard framework is decision-theoretic: consider an agent who has a menu of doxastic options (credal-states, belief-states, or plans to update them) and a purely epistemic utility function (standardly, a measure of accuracy). What option should they choose to optimally promote epistemic utility?

Epistemic consequentialism is a hot topic, for it offers a rare combination of (1) a fruitful, tractable framework that (appears to) provide insight into both old and new questions; and (2) apparently devastating problems that (threaten to) undermine any such insight. This combination has sustained a breakneck literature, with sympathizers claiming ever more insights, objectors claiming ever more problems, sympathizers offering ever more sophisticated responses to those problems, and so on.

On the side of the sympathizers, we *do* seem to learn something when we are shown a proof that—on natural ways of measuring accuracy—if a credence function is not probabilistic, it is guaranteed to be less accurate than some specifiable probabilistic one (Joyce 1998, 2009). At a first pass: since you should try to be accurate, you should be probabilistic!

But—on the side of the objectors—that first pass is clearly wrong. For suppose you are wandering through the garden of epistemic imps (Greaves 2013). Before you Imp 0 stands, clear as day: you're certain that she's outside. But she offers you an *epistemic bribe*: so long as you have any credence that Imp 0 is outside, exactly half of Imps 1–10 will come outside, and you will have no idea which ones; however, if you adopt credence 0 that Imp 0 is outside, all of Imps 1–10 will come out. If anything is epistemically irrational, it is adopting credence 0 that Imp 0 is outside: given your evidence, you should be certain that she is, and fifty-fifty that each of the other imps will be. Nevertheless, you *know* that if you adopt credence 0 that Imp 0 is outside and credence 1 that each of the other imps will be, then (since this will cause the other imps to come out) your doxastic state will be more accurate overall—you will trade certainty in one falsehood for certainty in many truths. Therefore, trying to be accurate can lead to clear irrationality.

Against this backdrop, the volume does what you'd expect. It combines a series of new applications of epistemic consequentialism, several new objections, and a handful of responses to old objections.

Applications first. Hilary Kornblith (chap. 3) argues that epistemic consequentialism vindicates a naturalistic epistemology. Alejandro Pérez-Carballo (chap. 6) shows how consequentialist tools can be used to evaluate courses of inquiry. Sophie Horowitz (chap. 11) uses such tools to show that “Jamesian permissivism”—the intuitive thought that different epistemic values can lead to different permissible responses to the evidence—is problematic. Amanda Askell (chap. 12) suggests that consequentialist tools can help formulate and problematize antiakrasia norms for credences. And Jeffrey Dunn (chap. 13) uses consequentialist tools to argue that epistemic tragedies of the commons are not only possible, but common.

Objections next. Clayton Littlejohn (chap. 1) argues that knowledge (not accuracy) is the fundamental epistemic value, and that knowledge's constitutive tie to rational belief prevents it from vindicating consequentialism. Nancy E. Snow (chap. 2) argues that "adaptive misbeliefs"—useful beliefs that are false—are problematic for epistemic consequentialism. Christopher Meacham (chap. 7) argues that accuracy-based arguments do not—and perhaps *could not*—vindicate standard norms like the Principal Principle (Lewis 1980). And Michael Caie (chap. 8) argues that attention to the details shows that thoroughgoing epistemic consequentialism completely crashes, offering no deontic distinctions between different doxastic states.

As this dialectic shows, the holy grail for epistemic consequentialists is to find an interpretation of their framework that preserves the (apparent) insight of the applications while avoiding the problematic consequences (see Carr 2017). Several chapters attempt to do just that. Julia Driver (chap. 5) argues that a distinction between wrongness and blameworthiness can help defuse the epistemic bribe objection. Richard Pettigrew (chap. 9) argues that the response to the bribe objection given by Konek and Levinstein (2019) fails, and instead defends an error theory for our intuitions. Ralph Wedgwood (chap. 4) argues that if we presuppose more about rationality, we can use consequentialist tools to account for epistemic correctness in way that is fruitful but doesn't sanction epistemic bribes.

James M. Joyce (chap. 10) offers a different reply. Slogan: "Credences are for *using*, not for *choosing*, and must be evaluated accordingly" (257). The problem with taking an epistemic bribe is that in doing so you choose a given credal state *not* because you want to use it to make estimates or to decide how to act, but simply because it'll bring about a good (accurate) state. Thus although adopting credence 0 that Imp 0 is outside would allow you to make your credences about the other imps more accurate, you would nevertheless prefer to make decisions as if you had credence 1 that Imp 0 is outside. Since credences are to be understood in terms of their functional role, taking the bribe involves adopting a "sham credence" 0. To make this precise, Joyce imposes a *ratifiability* constraint on credal states: it's rational to have a credal state *C* only if at some (hypothetical) prior time, *on the supposition that you'll adopt C*, the expected accuracy of adopting *C* is at least as great as the expected accuracy of any other option (258). The idea is that when this constraint is not met, upon adopting *C* you'd prefer to use a different credal state to guide your actions—thus *C* is not a genuine credal option at all. For instance, assuming that you'll adopt credence 0 that Imp 0 is outside, it maximizes expected accuracy (under that assumption) to switch to credence 1 that Imp 0 is outside—the original credal state is not ratifiable.

Although there is much to learn from Joyce's careful discussion, there's reason to worry about both the details and the big picture. Details first. On any reasonable account of credences, it's possible to have a given credence in *p* and not be certain of which credence in *p* you have. You can (for instance) be genuinely disposed to treat a bet on *p* as fair if and only if the bet is at 2:1 odds (i.e., to have credence $\frac{1}{3}$ in *p*), but be unwilling to bet the farm that you are so disposed (i.e., not have credence 1 *that* you have credence $\frac{1}{3}$ in *p*). That state—call it *C*—is a genuine credal state, not a "sham credence." Yet it fails Joyce's ratifiability test. For if you are in state *C*, you will be uncertain whether you are in state *C*. Thus *on the supposition that you adopt state C*, there is another state that is more expectedly accurate than *C*—namely, a state *C** that is certain

that you are in state *C* and is otherwise identical to *C*. That means *C* is not ratifiable. Thus ratifiability does not draw the right line between genuine and sham credences.

Turning to the big picture, Joyce's approach faces a motivational challenge. If credal states are *not* to be evaluated in and of themselves—but instead in terms of what they warrant *doing*—then why did we ever need a nonpragmatic vindication of probabilism (Joyce 1998)? The Dutch Book argument already shows that nonprobabilistic credences warrant acting in ways that lead to a sure loss. If—even in the context of pure epistemology—credences are for *using*, not for choosing, then why isn't that an indictment enough?

Those of us attracted to epistemic consequentialism should consider alternatives. What we need is *an* interpretation on which consequentialist tools are illuminating, but epistemic bribes are not sanctioned. Here—perhaps—is one. Call it *accuracy-first evidentialism*. It is a two-tiered theory. The first tier is simple: you are epistemically rational iff you conform to your evidence. That raises a question: what is the structure of evidence? And why does it have that structure? The second tier provides an answer: *evidence*—whatever else it is—is *an optimal guide to truth*. This means that it must have various structural features; enter the consequentialist arguments. Why is evidential support probabilistic? Because otherwise it'd be accuracy-dominated, and so wouldn't be an optimal guide to truth (Joyce 1998). Why does evidential support conform to the Principle of Indifference? Because otherwise it would be worst-case accuracy dominated (Pettigrew 2016), and so (perhaps) wouldn't be an optimal guide to truth. Why does evidential support warrant believing *p* iff it warrants having sufficiently high credence in it? Because otherwise the beliefs the evidence warrants wouldn't maximize expected accuracy by its own lights (Dorst 2019), so it wouldn't be an optimal guide to truth. And so on.

Accuracy-first evidentialism does not sanction epistemic bribes. Bribes happen when your *credence* is known to be causally or constitutively connected to various other claims, and therefore trying to maximize its accuracy leads you to change it in irrational ways. However, on this picture your job is *not* to try to have the most accurate credences—your job is to conform to the evidence. Since all agree that when you're staring at Imp 0 your evidence warrants credence 1 that she's outside, there is no pressure to have any other opinion.

In order to make a bribe on this picture, the evidential support relation itself would have to be (known to be) causally or constitutively connected to various other claims. I have no impossibility proof, but it is hard to see how this could be done. Plausibly, facts about what various bodies of evidence support cannot be causally connected to contingent facts about where someimps are. (Facts about which body of evidence you *have* could be so connected—but such facts are not relevant to the second tier, when the accuracy arguments are applied.) And absent liar-like self-reference, constitutive connections look dubious.

There are downsides to accuracy-first evidentialism. It does not apply as widely as the injunction “always maximize expected accuracy”; it does not attempt to explain the injunction to conform to your evidence; and it may face standard objections to two-tiered theories. Maybe it does not provide the grail the consequentialists are after. Maybe *no* interpretation will provide such a grail.

What would that mean, if so? Would the tools of epistemic consequentialism—and the debates of this book—fall by the wayside?

I think not. Whatever the role of accuracy in epistemology, it is at least *relevant* if your theory of rationality is guaranteed to lead to less accurate beliefs than some alternative. Consequentialist tools give a versatile way to explore such possibilities. Moreover, the versatility and precision of these tools allows them to easily extend to issues on the edge of epistemology, where traditional methods do not provide much traction. For example: How should social organizations and groups be structured, epistemically speaking? (Kitcher 1990; chap. 13 of the volume under review). When and why should you gather new information, and what sorts of information should you gather? (Greaves and Wallace 2006; Briggs and Pettigrew forthcoming; Campbell-Moore and Salow forthcoming; and chap. 6 of the volume under review). How should those who disagree make epistemic compromises? (Moss 2011). When should nonideal agents approximate ideal ones—and why? (De Bona and Staffel 2017, 2018). How should you update your beliefs when you might make mistakes? (Bronfman 2014; Gallow forthcoming). And so on. All of these questions—and many more besides—can be straightforwardly addressed using consequentialist tools in a way that is precise, tractable, and often yields surprising conclusions.

So? As a foundational theory of rationality, epistemic consequentialism may fail. But the *tools* it has generated will remain an important part of the epistemological landscape going forward. And we should be glad of it.

Kevin Dorst

Massachusetts Institute of Technology

References

- Briggs, R. A., and Richard Pettigrew. Forthcoming. “An Accuracy-Dominance Argument for Conditionalization” *Noûs*.
- Bronfman, Aaron. 2014. “Conditionalization and Not Knowing That One Knows.” *Erkenntnis* 79, no. 4: 871–92.
- Campbell-Moore, Catrin, and Bernhard Salow. Forthcoming. “Avoiding Risk and Avoiding Evidence.” *Australasian Journal of Philosophy*. Carr, Jennifer, 2017. “Epistemic Utility Theory and the Aim of Belief.” *Philosophy and Phenomenological Research* 95, no. 3: 511–34.
- De Bona, Glauber, and Julia Staffel. 2017. “Graded Incoherence for Accuracy-Firsters.” *Philosophy of Science* 84, no. 2: 189–213.
- De Bona, Glauber, and Julia Staffel. 2018. “Why be coherent?” *Analysis* 78, no. 3: 405–15.
- Dorst, Kevin. 2019. “Lockeans Maximize Expected Accuracy.” *Mind* 128, no. 509: 175–211.
- Gallow, J. Dmitri. Forthcoming. “Updating for Externalists.” *Noûs*. Greaves, Hilary, 2013. “Epistemic Decision Theory.” *Mind* 122, no. 488: 915–52.

- Greaves, Hilary, and David Wallace. 2006. "Justifying Conditionalisation: Conditionalisation Maximizes Expected Epistemic Utility." *Mind* 115, no. 459 : 607–32.
- Joyce, James M. 1998. "A Nonpragmatic Vindication of Probabilism." *Philosophy of Science* 65, no. 4: 575–603.
- Joyce, James M. 2009. "Accuracy and Coherence: Prospects for an Alethic Epistemology of Partial Belief." In *Degrees of Belief*, edited by Franz Huber and Christoph Schmidt-Petri, 263–97. Dordrecht: Springer.
- Kitcher, Philip. 1990. "The Division of Cognitive Labor." *The Journal of Philosophy* 87, no. 1: 5–22.
- Konek, Jason, and Benjamin A. Levinstein. 2019. "The Foundations of Epistemic Decision Theory." *Mind* 128, no. 509: 69–107.
- Lewis, David. 1980. "A Subjectivist's Guide to Objective Chance." In *Studies in Inductive Logic and Probability*, vol. 2, edited by Richard C. Jeffrey. Berkeley: University of California Press.
- Moss, Sarah. 2011. "Scoring Rules and Epistemic Compromise." *Mind* 120, no. 480: 1053–69.
- Pettigrew, Richard. 2016. *Accuracy and the Laws of Credence*. Oxford: Oxford University Press.