Philosophers' Imprint

volume 18, no. 12
june 2018

Epistemology from an Evaluativist Perspective

Hartry Field
New York University

1. Skepticism, and the plan of the paper.

There is a class of globally skeptical positions that many philosophers take somewhat seriously. Examples include skepticism about the external world, skepticism about inductive reasoning, and skepticism about logical inference. When I say that many philosophers take them somewhat seriously, I don't mean that they have any inclination to be globally skeptical in any of these ways, but just that they regard these forms of skepticism as posing a serious threat that demands a non-dismissive answer. But in my view, these skeptical positions pose a threat primarily to those who have been taken in by an erroneous meta-epistemology, a species of epistemological realism that I've elsewhere (Field 2009) derided as the "justificatory fluid" picture.[1]

The erroneous picture has it that epistemology is concerned with ascertaining *the facts about justification*, conceived as analogs of *the facts about the electromagnetic field*, or about *flow of electromagnetic energy*, or some such thing. On this picture, skepticism is the threat that justification in this metaphysical sense might not exist, or might have properties totally different than those our justificatory practices are thought to presuppose. And on that conception of justification, the threat of various forms of global skepticism does seem very serious.

Let's consider as an example the cartoon version of Hume's worry about induction, as given for instance in Wesley Salmon 1967. The "problem of induction", as Salmon poses it, is to discover the source of justification of our inductive practices. The fact that people engage in these practices is irrelevant: they might do so without the benefit of

1. I'm slightly overstating this, for at least in the case of skepticism about the external world, there might be positive arguments for thinking we go badly wrong: e.g. arguments, based on cosmological theories for which we have evidence, that we are more likely to be Boltzmann brains with a bleak future than the normally evolved brains in the kind of world we think we're in. (Somewhat similarly, a lab assistant in a lab with a thousand brains in vats who's told that these brains are all now being fed exactly her experiences might have reason to worry whether she is really the assistant — a worry that would presumably become more pressing were she told that in the near future the brains in vats would be decoupled from her and coupled instead to someone being hideously tortured.) The skeptical worries I mean to be dismissing are ones based not on any such positive arguments for skeptical hypotheses, but merely on the absence of arguments against them.

justification. The fact that we approve of these practices likewise cuts no ice: we might be wrong in so doing. To say that the practices *work* would be question-begging. Sure, they've worked well in the past, but to conclude from this that they will continue to do so in the future is to make an inductive argument; if there isn't already a justificatory source for inductive reasoning, this will cut no ice, and if *real justification* is counterinductive, it will make things worse.

The best hope, Salmon tells us, is for a pragmatic justification (or "vindication") which tries to argue that our inductive methods are better than any competing methods. But he thinks that the argument for their superiority cannot rely on empirical conclusions established by induction: that would be question-begging. What's required instead, he thinks, is a mathematical proof that our inductive methods, and they alone, have a certain combination of desirable properties; without that, it's hard to see why we should prefer science to voodoo (1967, p. 55). But as he came to realize, it is very difficult to find clearly desirable properties that our methods can be mathematically proved to have, beyond very weak properties that clearly undesirable methods have as well. (Moreover, if the skepticism under consideration were extended to include logical and mathematical reasoning, we'd be in even more trouble.)

Here's my cartoon of the cartoon: In Salmon's view, what's needed for induction to be reasonable is that premises of an inductive argument pass justificatory fluid to the conclusion. Given such a picture, it seems correct that without a positive argument for such justification we shouldn't believe in it: our default position should be the "skeptical" one that there is no such fluid. And it's hard to see what that positive argument for the "justificatory fluid" could be. Similarly, it's hard to see what positive argument there could be that, if it exists, it flows in the "inductive direction" as opposed to the counterinductive.

Arguments from the regress of justification have some force on an epistemological realist picture: they make vivid (what should be plausible even without them) that you don't create "justificatory fluid" *ex nihilo*. Of course this could be denied, and coherence theorists do: build enough connecting pipes and the fluid will appear to fill them. But that gives rise to the standard objection, which seems pretty decisive: lots of *coherent* methodologies for reasoning from the observed to the unobserved are obviously *bad*.

A realist could instead hold that we needn't create the fluid *ex nihilo*, since it is there from the start: we have an "*a priori* entitlement" to believe certain things (such as that we aren't brains in vats) without evidence. This indeed, if developed in a certain way, may lead to something close to the methodology that I will recommend, but the nature of these "entitlements" is *prima facie* quite mysterious. Is it that while God gives us black marks if we believe without evidence that Hillary Clinton ran a child porn ring from a pizza parlor, he doesn't give us black marks for believing without evidence that we aren't brains in vats? (Or maybe it's just like that except without the God?) Also, wouldn't there be room for doubts about whether the "entitlements" really favor our practice over practices that we regard as bad ones, and shouldn't that meta-skepticism tend to diminish the entitlement?

Perhaps these remarks are built on too loaded an interpretation of 'entitlement'; indeed, perhaps the anti-realist view to be sketched supplies an anodyne interpretation of the nature of "entitlements" that could be regarded as vindicating the entitlement view.[2] But I think that without such a questioning of epistemological realism, it will be hard to set skepticism to rest.[3]

I concede that the label 'epistemological realism' is less than totally clear. I also concede that there are views that naturally fit under this label for which the "fluid" metaphor is inappropriate, and which the previous paragraphs don't adequately target. I'll consider one such

2. For a sophisticated version of the entitlement view, see Wright 2014.

3. I don't want to quibble about the word 'skepticism'. Indeed, in one sense of the term, just about everyone today is a skeptic: we've all gotten beyond the Cartesian hope for an *a priori* and incontrovertible proof that our inductive methods couldn't possibly lead us badly astray. The main target of these opening remarks is those who don't believe in the possibility of such a proof, but do believe in (or hanker after) some kind of "metaphysical justification" short of that.

form of realism in Section 2, but argue that it too fails to handle skepticism properly. But I'm less interested in coming up with a general argument against "epistemological realism" (however exactly that might be defined) than in sketching an alternative view of epistemology that I think does not generate problems (of which skepticism is only one) to which standard ways of thinking about epistemology give rise.

I'll sketch the alternative starting in Section 3. One aspect of it involves (to put it very roughly) focusing on sensible epistemic practices without fetishizing such notions as justification and knowledge. There's a lot of philosophical baggage currently built into these notions, which can to some extent be avoided by framing epistemological questions in such terms as "What would it be reasonable to believe in this situation?", "How confident should I be?", and so forth. Of course there's no reason to banish ordinary epistemological terms like 'justified' and 'knows': for instance, 'justified' can be used to mean 'reasonable'. But then it's unobvious that it makes sense to talk about "the nature of justification", and the idea that for a belief to be justified there must be a "source" of "the justification" loses much of its force.

Such an alternative to the sort of epistemology that takes skepticism seriously does not put basic features of our inductive practices beyond debate: there can be serious questions about which of our inductive practices need improvement, and about the ways in which they should be improved. But debates about these matters are best conducted in the spirit of the sailors fixing Neurath's boat while it's at sea, so that only *specific local objections* should worry us about extant inductive practices. Section 3 will sketch what an epistemology more focused on such local improvements in basic methods might look like, broadly in keeping with some remarks of Reichenbach. Section 4 deals with an objection that might seem to undermine the coherence of looking for local improvements in this Reichenbachian spirit.

I don't say that it's impossible to combine the "Neurath's boat" methodology with an epistemological realism — some "entitlement" views, e.g. Wright 2014, seem to do so — but I do think that the epistemological realism makes such a picture awkward. After all, there are possible "boats" that we think totally unacceptable. Presumably someone who is on such an unacceptable boat shouldn't count as justified, however well he or she does at meeting local objections. But if "facts about justification" are conceived of in a realist spirit, it seems like mere dogma to assume that the boat we're floating on is one of the "justified" ones. This is likely to lead to skeptical doubts at the meta-level: doubts about whether one is justified. And doubts about whether one is justified have some tendency to induce skeptical doubts at the ground level.

Ultimately, then, there is a tension between any Neurath's boat methodology and epistemological realist positions: the epistemological realism tends to undermine the methodology. (Shifting from somewhat loaded epistemological notions like "justification" to blander ones like "reasonable" helps only in that an epistemologically realist understanding of the latter is less likely.) I won't try to argue that there is no way around this, but my preferred solution is to give up on the realism and offer a different understanding of notions like being justified or being reasonable. One way to understand them which I think would disarm the skepticism is a blatantly subjectivist one: for a belief or method to be justified or reasonable is just for it to accord with the standards of the believer or the believer's community. This, however, strikes me as a highly implausible line. It's no better to replace 'the believer's community' with 'our community' (or with 'an assessor's community')[4]: this would preserve the fundamental problem with subjectivism, which is its attempt to explain normative terms as descriptive. My preference is for something more along the lines of Gibbardian expressivism (or MacFarlane-style assessor relativism, *properly interpreted*: see preceding note). I will not enter into the details of how such a view is to be formalized, but will make some general

4. MacFarlane 2005 is sometimes interpreted as making the latter proposal, and some of his remarks do suggest it, but it seems to me an incoherent view. I think his paper suggests a far more interesting view, very much akin to Gibbardian expressivism, and indeed he has recently said (2014) that the difference between his view and Gibbard's is rather slight. The difference he notes won't matter for purposes of this paper.

remarks in Section 5 aimed at those unsympathetic to any such view and at those who think it couldn't help with the skepticism.

The paper, then, presents a package that combines a methodology for epistemology (that has antecedents in Neurath and Reichenbach, among many others) with a normative anti-realism. The two parts of the package are to some extent separable. As I've already said, the Neurath's boat methodology may be compatible with a more realist view of epistemology, though there is a tension between them. Conversely, the anti-realist view of epistemology that I will recommend doesn't require the methodology: for instance, it is probably compatible with a more foundationalist response to skepticism. Indeed, one could make further assumptions compatible with the anti-realism that would reinstate some of the force of the kind of skeptical arguments from which we began. I'll discuss that in Section 6. Nonetheless, I do think the methodology I'm proposing is very natural given the anti-realism, much more natural than without it, and for that reason I think it reasonable to present the two parts of the package together. (I'll use 'evaluativism' as a name of the package, but it will generally be clear in any context whether it is the methodological or metaphysical aspect I have in mind, and I will sometimes be explicit.)

2. Reliabilism and concept constitution.

The "fluid" metaphor is most naturally associated with non-naturalist forms of epistemological realism. *One* alternative to that is a kind of reliabilism; the idea is to forgo the mysterious fluid in terms of a perfectly naturalistic property of reliability.

I don't think this is the way to go. In part that's for a rather standard reason — that reliabilism doesn't seem to do justice to the internalist aspects of epistemology — but I will defer that criticism, since inductive skepticism isn't the best place to raise it.

There's another reason that I think more decisive in the inductive case: there's simply no notion of reliability adequate to the job. One feature of inductive methods is their "self-correcting" character: if these methods are applied in circumstances in which they initially behave unreliably, they typically start to behave more reliably. This is so for many bad inductive methods as well as for good ones; it isn't easy to see how to come up with a definition of reliability according to which, even among "equally powerful" rules, the good ones are the ones that are "most reliable".[5] There is a wide variety of externalist good-making features in an inductive method; I have no doubt that one method scoring higher than another with respect to most of these external good-making features contributes to it being the better method, but I think it extremely unlikely that there is any way of capturing this with any simple notion like reliability. (And as hinted, such externalist factors do not exhaust the relevant considerations.)

In the next section, I'll present my favored alternative to both justificatory fluid approaches and reliabilist approaches. But first let's turn to another kind of skepticism: skepticism about logical knowledge. Lewis Carroll (1895) gives a classic presentation of one problem here, the problem of how we can know anything via logical inference. The problem, as he presents it, is especially salient on a justificatory fluid perspective: how do the justifications for A and for $A \rightarrow B$ combine to squirt justification into B? (If you prefer: How do they "transmit warrant to" B?) He considers the answer that it's because we have justification for the premise $A \wedge (A \rightarrow B) \rightarrow B$, but he then points out that that doesn't seem to help unless A, $A \rightarrow B$, and $A \wedge (A \rightarrow B) \rightarrow B$ squirt juice into B, which itself relies on the assumption that Modus Ponens is juice-squirting.[6]

One way around the Lewis Carroll problem is to go reliabilist: we don't need an internal justification of logic, for logic is "justified" in the only sense that matters as long as it's reliable. And reliabilism in this case is more promising than in the inductive case, because in this case

5. There's a bit of discussion of this in Field 2000, Section 4.

6. There's an irony in trying to justify Modus Ponens in terms of the validity of the schema $A \wedge (A \rightarrow B) \rightarrow B$, technically known as "Pseudo Modus Ponens": a moral of the Curry paradox, at least for those who accept restrictions on classical logic to ensure naive assumptions about truth, is that Modus Ponens is acceptable without restriction only if Pseudo Modus Ponens is *not* acceptable without restriction.

we have a pretty clear account of reliability: a form of logical inference can be regarded as reliable if, of logical necessity, it preserves truth.

The main problem (or at any rate the main non-technical problem)[7] for reliability here is the one I deferred in the case of induction: exclusive focus on reliability doesn't do justice to the internalist aspects of epistemology.

Paul Boghossian (2003) has emphasized one important way that it doesn't: according to the obvious version of reliabilism, inference via a highly unobvious deductive rule that can be shown reliable only by an extraordinarily complicated mathematical proof far beyond a person's grasp would "justify" the person's conclusions, whereas intuitively such a person is making totally unjustified logical leaps. Maybe some sort of "higher-order" reliability considerations could be invoked in the hope of getting around this, but it is far from obvious how that would go, and some of the worries about the clarity of the notion of reliability in the inductive case would then be likely to arise in the deductive case as well.

For another way to see how unintuitive the thoroughgoing externalism about the epistemology of logic is, consider debates about logic. Let's suppose that Hilary, Michael, and Saul disagree about logic: Saul believes in classical logic, Michael in intuitionist logic, Hilary in quantum logic. And let's pretend that each has come up with a well-worked-out view according to which his favored logic is the correct one, with strong *prima facie* arguments favoring his logic over the other logics. (Each has a *prima facie* reasonable reply to the others' arguments, so none of the arguments is clearly decisive; it's the kind of situation where the evaluation of the argument might turn on very high-level theoretical considerations of, say, the role of logic.) Now consider a typical logical inference that is valid according to one of these logics but not according to another: say one from a premise of form "not both A and B" to the corresponding conclusion "either not A or not B" (which is not intuitionistically valid, but is valid in quantum logic as well as classically). According to the reliabilist, whether one of our three characters is justified in inferring in this way depends not a whit on their logical views: if this De Morgan law is not in fact correct, then none of the three would be justified in making the inference, whereas if one of the other logics is in fact correct, all three are; *and that is all that can be said about the matter*.

I'm not at all opposed to the idea that our epistemic approval or disapproval of someone's reasoning might be to some extent conditional on the "external" question of which logic is actually correct,[8] but I find it hard to believe that actual correctness is the whole story: if in some context Michael slips and reasons as above even though that reasoning doesn't accord with the logic he's advocating, then his reasoning has a problem that Hilary's and Saul's reasoning doesn't have when they reason in that way; and this is so whichever logic is correct.

The case can be made more decisive by imagining that the entire logical community unanimously supports, by arguments that we all now find compelling, a logic that is not actually correct (and maybe will someday be shown incorrect by revolutionary geniuses); shouldn't the overwhelming theoretical support for what will later be seen as an incorrect logic count for *something*? Again, I'm not denying that the correctness of the logic might be one factor in the evaluation of their reasoning; I'm just denying that it is the only factor.[9]

7. A technical problem is that when we extend to logics of truth, we can no longer equate good logics with ones that preserve truth by logical necessity. There's a discussion of this in Field 2015.

8. Let's put aside any doubts one might have about the presupposition that one logic is uniquely correct. (Even on the assumption that correctness for logic is truth-preservation by logical necessity, this could be questioned on the grounds that there needn't be a unique notion of logical necessity.)

9. I think the same kind of argument can be given in the inductive case too, though for it to have much intuitive force we need to restrict it to cases where the alternative methods are within the spectrum of reasonability: otherwise the "external" aspects of reasonability overwhelm the phenomena. Consider two people, one of whom is generally more cautious than the other about inferences to the unobserved, though the first isn't crazily cautious and the second doesn't crazily jump to conclusions. If on a given occasion the first person goes against his usual methodology by making an inference to the unobserved that is warranted only on the other's methodology, isn't there an important sense in which he is less justified than the other person in making

Boghossian himself offers an alternative to reliabilism within what I'm calling the "juice" framework (Boghossian 2003), in which the "juice" is supplied by the meanings of concepts. His paper is framed around the question of why the premises of a Modus Ponens "transmit justification" to its conclusion. His answer (or, at any rate, the central part of it: see note 12 below) is that it's because reasoning in accordance with Modus Ponens is a precondition for having the concept *if … then*, which is an ingredient in the rule.

A complication in discussing this is that there is more than one concept *if … then*. As is well-known, the "material conditional" $A \supset B$ (defined as $\neg A \vee B$) is not a good account of the ordinary English 'if … then' — witness 'If I run for President in 2020, I'll win', which is true on the $\supset$ account (and not because I'll win). On the other hand, the conditional » employed in such examples is completely inappropriate for another task of the conditional, restricting universal quantification: it may well be true that everyone who will be nominated by a major party for the 2020 election is female, but it certainly isn't true that ∀x(x will be nominated by a major party for the 2020 election » x is female), since that implies 'Ted Cruz will be nominated by a major party for the 2020 election » Ted Cruz is female', which is false on the above stipulation for » even though its analog for $\supset$ is probably true. The point of this is just that in discussing Boghossian's claim, we need to decide whether we're talking about the role that Modus Ponens for $\supset$ plays in the meaning of $\supset$, or the role that Modus Ponens for » plays in the meaning of ». The cases are structurally similar, but different in detail.

Part of their similarity is that in both cases, there's at least one prominent view (which is well-motivated even if not ultimately compelling) that denies Modus Ponens.

- In the case of $\supset$, the prominent view is dialetheism, the view that under certain circumstances it's allowable to simultaneously accept both a sentence and its negation, but where the damage typically associated with accepting such contradictions is limited because they don't imply everything. It is almost immediate that dialetheists must reject Modus Ponens for $\supset$, if $\supset$ is defined as above in terms of $\neg$ and $\vee$: if one accepts both A and $\neg A$, one will surely accept both A and $\neg A \vee B$ no matter how absurd the B, so Modus Ponens for $\supset$ would require one to accept absurdities.

- In the case of the ordinary conditional », one such prominent view is McGee's (1985; see also Kratzer 2012 Chapter 4), according to which » obeys the Exportation Principle

$$\frac{(A \wedge B) \gg C}{A \gg (B \gg C).}$$

That, with Modus Ponens for », leads to the rule

$$\frac{A \qquad (A \wedge B) \gg C}{B \gg C,}$$

that inference? And isn't that so independent of which inductive method is "correct", if talk of correctness here even makes sense?

> to which there are clear counterexamples,[10] and so McGee rejects Modus Ponens for ».[11]

Williamson 2007 (focusing on the McGee case, but the point generalizes) has pressed the claim that reasoning in accordance with Modus Ponens for a conditional can't be a precondition for having the concept of that conditional, because prominent and well-motivated views keep the concept while rejecting the alleged precondition.

One obvious way around the letter of Williamson's critique, and a way that Boghossian takes, is to reject the view that the dialetheist has the same concept of ⊃ that non-dialetheists do, and that McGee has the same concept of the ordinary English 'if … then' as those of us who keep Modus Ponens for it but reject Exportation. But I don't think this ultimately helps. Though Williamson himself takes a strong stand on the issue of when there has been a change of concept, he needn't. The basic point is: maybe it will someday be shown that reasoning with the standardly accepted rules for ⊃ and/or » leads us astray in some circumstances. If so, we will want to reason using different rules, for connectives that we can call ⊃* and »* (that we may or may not regard as "the same concepts as" ⊃ and »). If we want to say that the concepts ⊃* and »* differ from the concepts ⊃ and », fine: in that case, the concepts ⊃ and » are bad ones that will lead us astray. But in that case (at least once we've *seen* that they lead us astray), we can't regard reasoning in accordance with the rules for ⊃ and » as legitimate, despite the rules being meaning- constituting. Meaning doesn't have the epistemological clout that Boghossian requires.[12]

Incidentally, even if the meaning line worked for the logic case, it seems hard to apply in a remotely attractive way to the inductive case, since inductive rules don't involve any special connectives.[13] Pollock (1987, sec. 4) did try to generalize it to that case, by proposing that *every* empirical concept a person possesses is so shaped by that person's system of epistemological rules that there can be no genuine conflict between the beliefs of people with different such systems; as a result, the systems themselves cannot be regarded as in conflict. But this view is wholly implausible. I grant that there's *a sense* in which someone with even slightly different inductive rules inevitably has slightly different concepts of *raven* and *black* than I have, but it is not a sense that licenses us to say that his belief 'The next raven will be black' doesn't conflict with my belief 'The next raven will not be black'. It seems hard to deny that there would be a conflict between these raven beliefs, and if so, the systems of rules give genuinely conflicting instructions.

In any case, the point from the logic case remains: declaring certain inductive rules "concept-constituting" does nothing to show that they can't be legitimately criticized; it just stipulates that the criticism will be regarded as a criticism of the concepts (here, *raven*, *blackness*, and all other empirical concepts). The old inductive rules are de-legitimated by the criticism (if the criticism is good), whatever one's view about whether the concepts have changed.[14]

10. If you don't like McGee's election example, let A and C both be 'I'll eat dinner tonight' and B be 'I'll be beheaded a moment from now'. (But I think this example makes it pretty clear that it is Exportation rather than Modus Ponens which is problematic.)

    Kratzer 2012 argues, fairly convincingly, that it's a mistake to represent ordinary conditionals in terms of a primitive operator »: instead, those without an overt modality have the form Must(q|p) where this is in effect a binary operator ("on the assumption that p, it must be that q"). She takes it that when » is so defined, then McGee is right that Modus Ponens rather than Exportation is to blame. But this last is far from obvious, it depends on her interpretation of "stacked relative clauses" (p. 105) according to which Must(Must(p|q)|p) is equivalent to Must(p| p∧q) and thus trivial, which seems surprising given that $\Box(p\supset\Box(q\supset p))$ is invalid in virtually every modal logic.

11. Another view, perhaps related and probably more defensible, is that of Kolodny and MacFarlane (2010), on which » violates Modus Ponens for some sentences with deontics or epistemic modals in the consequent.

12. Actually Boghossian does allow that the meaning-constituting rules of some concepts make those concepts defective, and that that undermines any justification that their meaning might provide. But he rules out this happening for conditionals — and, I think, for other logical concepts except for *transparently* defective ones like *tonk*. He doesn't seem to allow for cases where there is a serious theoretical issue as to whether a given logical concept is defective.

13. This has been noted by others, e.g. Enoch and Schechter 2006.

14. Pollock's view is that it is our object-level concepts like *raven* that are determined by our system of rules. A *slightly* more plausible view is that our

3. Systematic epistemology.

It's time to sketch out a different perspective on skeptical problems of the sort we've been considering. Let's start from the idea that the point of epistemology is to evaluate our own and others' methods of forming and retaining beliefs, typically in order to influence them to improve those methods (or to resist changes in the methods that would make them worse).

One feature of typical evaluations is that they are multi-faceted: "The movie presents a compelling situation and has imaginative cinematography, and the lead actress gives a knockout performance, but an important subsidiary character is poorly developed, and there is a hole in the plot." Epistemological evaluations are often like that too: "His conclusion was based on a good though unpopular method of statistical inference — ironic, since he actually advocates a different method which would have led to a different conclusion — and he makes good use of very extensive data, though there is a slight bias in the method by which he collected that data, and there is other available data that, if not accounted for, would seem to undermine his conclusion." It seems almost as absurd to evaluate beliefs on a single scale of degree of justifiedness as it is to evaluate movies or pieces of music or literature on a single scale of degrees of goodness. I don't mean to suggest that an epistemological realist would have to disagree with this; but I do think that there is a strong tendency in the realist literature to talk as if there were a single scale of justification.

Another feature of typical evaluations is that we make them using our own beliefs and preferences. To some degree we are willing to back off from these beliefs and preferences: "I can't stand Frank Sinatra, though I can see how, if you're into that sort of thing, he's pretty good at it." But there is little point in trying to back off to a position of complete neutrality.

This is of immediate relevance to the sort of skeptical arguments we've been considering. From an "evaluativist" perspective it is hard to see the point of the foundationalist demand for non-circular "justifications" of our inductive methods: there is no reason to think that the legitimacy of inductive arguments, or arguments by Modus Ponens, is in peril unless they can be non-circularly "grounded" — say, in the meanings of component terms.[15] To the question "Why use our inductive methods rather than counterinductive methods?" or "Why use Modus Ponens rather than affirming the consequent?" it seems perfectly fine to give the obvious answer, "Those other methods would yield radically wrong results", and this answer needs no further defense. Part of the reason this is fine is that we can back off quite a bit from the details of our methods without compromising the answer. That is, *any alternative inductive or deductive method that we can take remotely seriously* will agree with ours that counterinduction or affirming the consequent leads to absurd conclusions. Evaluating our methods as better than those doesn't require a completely neutral standpoint. That's in part because evaluating them in this way doesn't involve claiming that beliefs arrived at by our methods have something called "justification" for which we can sensibly ask, "From where does it flow?"

This might suggest either (I) that no issue of justification can arise for our most central deductive or inductive methods, or (II) that there is no role for a systematic epistemology. But I emphatically reject both views.

Regarding (I), I think that justification of central presuppositions is important in a broadly dialectical context. Suppose someone challenges current standards of deduction and induction by offering alternative standards that she regards as superior, or at least thinks might be superior. (The someone might even be ourself: 'dialectical' isn't

epistemological concepts like *reasonable* are so determined: 'reasonable' just means 'reasonable according to our (the assessor's) rules'. That modified view doesn't seem attractive either, but in any case, it wouldn't serve Pollock's purposes. For the advocates of alternative systems of rules would still be in genuine conflict about ravens, and each could raise skeptical worries about whether it mightn't be better to shift from the system that is reasonable in their own sense (viz., their own system) to the system that is reasonable in the other person's sense (viz., the other's system). All that the modified view would do is strip away the normative aspect of the term 'reasonable'.

15. To some extent it is desirable to systematize our evaluations, but there's no obvious reason why systematization needs to take a foundationalist form.

intended to exclude debates with oneself.) Then, to the extent that her considerations move us even though they don't ultimately convince us, we need to consider what can be said for why our standards are better than the alternative she's suggesting. There are some big issues here, but they will come up more clearly if I first turn to my alternative to (II): the nature of a systematic epistemology.

One important role for epistemology is the development of formal models of ideal epistemic behavior. The models that have been developed so far are extraordinarily oversimplified — for instance, Bayesian models don't handle failure of logical omniscience or even the invention of new theories, they involve superhuman computational complexity, they treat the notion of "basic observation propositions" as a black box, and attempts to say which priors are good are hopelessly limited (e.g. the continua of inductive methods are confined to languages with only monadic predicates). Despite such extraordinary limitations, Bayesian methods are extremely illuminating for their resolutions of a wide range of puzzles, and it is important to try to develop far more realistic models that incorporate their insights.

What are these models models of? One thing to model is how people actually do things. I'm talking here about an idealized model, one which abstracts away from mistakes due to tiredness, inattention, drunkenness, and so on. (There may be some dispute as to which features of our performance are mere "performance errors" that should be idealized away, but there is no reason to insist on a hard and fast decision in all cases: one can look for models that idealize the feature away and models that build it in.) Let's call the task of coming up with such models the *quasi-descriptive task* of epistemology. (The terminology somewhat echoes the opening section of Reichenbach 1938.)

We can see, even from the crude models we now have, that a quasi-descriptive model of one person is unlikely to be the same in every detail as a quasi-descriptive model of someone else: for instance, in the continua of inductive methods there are one or more parameters that determine various features of caution about how to modify predictions about future instances, or belief in universal generalizations, on the basis of evidence. In more serious models there is likely to be far more opportunity for such variations. Even for a single person, it's hard to believe that there's a uniquely best choice of all such parameters for an idealized model of that person; and (more to the present point) it's hard to believe that the range of best choices for the idealized description of one person will be the same as the range for the idealized description of everyone else.

In any case, there is no reason to restrict our epistemological task to those that are quasi-descriptive of actual beings: it's also possible and I think important to invent and study methods without regard to whether anyone actually employs them. Maybe such methods would be better. The detailed formulation of such methods can be thought of as quasi-descriptive in an expanded sense: it quasi-describes possible beings (who might also be subject to tiredness, inattention, and drunkenness from which the quasi-description abstracts).

This brings us to the evaluation of methods: the "critical" and "advisory" tasks of epistemology, in Reichenbach's phrase. Here we study the different methods that we've isolated in the (expanded) quasi-descriptive phase, see how they perform in various circumstances, and make an evaluation of them based on this. Presumably the evaluation is to be comparative: we need to compare each method to other available methods, since in some sense we can't do better than using the best *available* method. Of course, we can try to make new methods available, and if we think that the best available isn't good enough, we will be motivated to try to do so; our degree of optimism about finding a better one might affect the degree to which we call the best available one deficient. We might also in some circumstances make a bet that a better method than we currently have will yield a certain verdict, and go with that verdict despite not having a very good backing for it. There is no formula for how to do all this; but once we give up the idea of a single scale of evaluation, there is no need for one. Trying to decide exactly what is required to be "justified" distorts good epistemological practice.

In Reichenbach's own practice, the evaluation of methods was supposed to be *a priori* and from a completely neutral standpoint: he was the initiator of the attempt at an *a priori* and non-circular "pragmatic vindication" of induction of the sort I've mentioned in connection with Salmon. But this part of Reichenbach's story can be separated from the rest, and I think we should drop it. What we want is a method that will work well *in a world like ours*, and our only hold on the features of a world like ours is through our inductive methods. We can back off a bit from our own beliefs and standards in making the evaluation of methods, but the restriction to a completely neutral standpoint is hopeless.

4. Defusing an objection to the methodology.

There is a worry one might have about this "evaluativist" methodology (indeed, I confess to having taken the worry too seriously in the past): that the critical/advisory task won't, in the end, cut any ice, because each method that emerges in the (expanded) quasi-descriptive phase will end up recommending itself.

That is the worry that emerges from David Lewis's two papers on "immodest inductive methods" (1971, 1974). Lewis considered as sample methods the methods of Carnap's continuum. In the first paper, he argued that the only method that recommended itself was an obviously inadequate one. In the second, he observed that there was a technical error in the first paper, which when corrected showed the problem to be, in a sense, even worse: every method in the Carnapian continuum declared itself superior to every other such method, so that self-evaluation simply has no force.

Lewis's conclusions are based on controversial rules by which a method scores itself and other methods, but I will not object on that count, since it is hard to find alternative scoring rules that lead to satisfactory results. My objection, rather, is to the significance given to these immodesty arguments.

Consider two kinds of proposal that seem rather analogous to proposals to revise our inductive methods.

First, proposals to revise *deductive* methods. Such proposals have been made for various reasons, some much better than others: the better ones include dealing with vague predicates in a way that resists arguments that such predicates have sharp boundaries, and dealing with truth and related notions in a way that allows naive principles to hold without leading to paradox. Advocates of such proposals present various reasons for adopting them, in the hope of persuading those who advocate the use of classical logic everywhere (even for vague predicates and/or for truth in paradoxical situations) to change their mind. If they do their job well, the reasons they provide won't depend on their preferred logic: that is, the arguments that they give won't use any logical principles that they disagree with the classicist about.[16] For we can typically show, in a background logic neutral between the two in question, that accepting one logic leaves such and such possibilities for vagueness and truth while accepting the other logic leaves so and so other possibilities; and then the advantages and disadvantages of each can be assessed, again in a way that is argumentatively neutral. (By "argumentatively neutral" I don't of course mean that its arguments are neutral between all logics, which would be impossible; I mean that they are neutral between the logics currently being debated between.) Doubtless, advocates of the different logics will be initially inclined to weigh the advantages and disadvantages differently; still, novel arguments for the overall advantages of logic L* over logic L may eventually persuade the advocate of L to try to modify her modes of reasoning.

16. Of course, if the preferred logic is in every respect weaker than classical, the neutral logic will just be that weaker logic. But normally, an alternative to classical logic is weaker in the nonmodal claims it accepts, but stronger in its rejections and perhaps its modal acceptances: e.g. if it is weaker in not accepting all instances of excluded middle, it will be stronger in *rejecting* some such instances (as opposed to accepting their negations), and in some cases may accept that certain negations of excluded middle are at least possible.

Moreover, even were one logic strictly weaker than the other, this wouldn't prevent rational debate: for instance, the advocate of the weaker logic can make the case that certain commitments of the stronger logic are uncomfortable, and the advocate of the stronger logic can make the case that the weaker logic is cumbersome.

But wait, doesn't each party in a debate about logic have an easy answer that settles the debate? E.g. can't a classical logician respond to any position whose coherence depends on a rejection of excluded middle, just by using excluded middle to show that position to be incoherent? ("The position is logically inconsistent, which is as bad as a position gets!") On the other side, can't an advocate of naive truth theory respond to any truth theory in classical logic by using naive truth to show that classical principles lead to absurdity?

The answer is that of course such question-begging arguments are available, but that in the context in question, they have little dialectical force. There might be value in stating such arguments if doing so makes clearer how each party views the other's position; but one shouldn't view such arguments as trumping other considerations, presented in a logic neutral between those under consideration, that might cut in the other direction.

Another example with the same moral concerns observational practice. Consider "Feyerabend cases" (Feyerabend 1975), cases where the old observational practice is laden with a theory that can be questioned. Feyerabend's own example concerned observations of the paths of falling objects. Feyerabend insists that when pre-Copernicans reported objects as falling in straight lines, they didn't mean "straight relative to the observer"; they meant "absolutely straight". Let's play along with that — we can imagine an alternate history in which it would be plausible, and his point doesn't really depend on historical accuracy.

We could imagine a dogmatic pre-Copernican using the old observational practice to dismiss the Copernican theory: "We've observed thousands of bodies falling in straight lines, whereas Copernican theory says they fall in curved arcs because of the spinning Earth; so Copernican theory is decisively empirically refuted!" But obviously that would be an absurd methodology: the right methodology is to develop the old theory and the new theory as best one can, and try to compare them in as neutral a way as possible, which in this case requires deploying an observational vocabulary (the language of relative motion) that is neutral between the theories at issue.

Returning to the case of induction, the point is that immodesty arguments, where e.g. each method declares itself best, are analogous to the ham-handed arguments against conceptual revision discussed above in the deductive and Copernican cases. Yes (putting aside qualms about the scoring measures they rely on), the advocate of a particular inductive method can argue, using that method, that that method is best. But as in the deductive and perceptual cases, such arguments do not preclude alternative arguments for the opposite conclusion. And there is no reason why the first argument should trump the second.

I wouldn't be happy to state this criticism by saying that because the first argument is circular, it has no force. Arguments that are in some sense circular can sometimes have a certain kind of force: for one thing, they can serve to illuminate what the position being argued for has to say about alternatives. An argument that is circular in this way can be especially useful when (as in response to most brain-in-vat scenarios) the position it begs the question against has nothing positive going for it. My claim is only that when the alternative does have a lot going for it, it is good practice to take the alternative seriously: to try to give a comparative evaluation of the two positions that is as neutral as possible.

In all three kinds of fundamental conceptual change (deductive, perceptual, and inductive), we have arguments for competing conclusions as to which alternative is better. So different aspects of one's theoretical state are pushing us in incompatible directions, and until we've resolved which direction to go in, we're in an incoherent belief state. I have given some vague advice about how to go about reasoning in such cases: "Develop each of the alternatives, even if they conflict with prior observational practice, logic, or inductive practices. See which does best." But if one wants to develop something more precise in these cases of conceptual change, we'll need a model of how to deal with inconsistent or otherwise incoherent belief states.

Such a model was suggested by Bryson Brown and Graham Priest (2004). They were dealing with classically inconsistent theories in domains where the appropriateness of classical reasoning is not in doubt[17] (such as the theory of infinitesimals that Berkeley critiqued, or Bohr's early model of the atom). Their general idea:

- Our cognitive processes are divided into "chunks", within which we reason using our logic.
- Instead of allowing free passage of information between chunks, we impose restrictions.

(E.g. in the case of infinitesimals, the first chunk assumes that infinitesimals are non-zero, and uses this to derive such conclusions as that the derivative of the function x^2 is the function $2x + dx$. This chunk passes that conclusion [though not the derivation] to the second chunk, which contains the premise that infinitessimals are zero and so concludes that the derivative is $2x$.) This model needs to be generalized a bit if it is to be applied to revision of fundamental practices (e.g. deductive, observational, or inductive). For this, we probably want an indeterministic model of mentality, where something like chance plays a role in both

- what theories (e.g. logical theories) one thinks up, and
- what evaluation one comes to of the respective merits of the theories.

Presumably focusing on one chunk and becoming influenced by its conclusions will diminish the influence of incompatible chunks; so we don't want the *fixed* restrictions on information transfer assumed by Brown and Priest.

But whatever the details, there is little doubt that we have rational ways of dealing with inconsistent premises. And we can then use this way of dealing with inconsistency, in the case where we have good arguments for substantial change of logic, inductive methodology, or observational practices, competing with the obvious ("question-begging") arguments against such change.

17. So the 'paraconsistent' in their title is potentially misleading.

Objection: a mental model of how we deal with inconsistency would tell us how we do change logic, not how we should. It would leave open the question of whether acting in this way is rational.

Reply: Once we have a model of how we do change logic (or even of how we might), the question of whether the model makes the change rational is simply a question of evaluation:

> Is acting in accordance with the model a good thing (or would it be, if we don't actually act that way)?

To answer this, we must compare the model to alternative models. And it's hard to believe that the dogmatic models could win.

In summary, the key features of rational revision of inductive methodology are:

- coming up with an alternative methodology (described in enough detail),
- arguing for merits of new methodology over the old (using the old, or what's common between the two), and
- retraining ourselves to operate in accordance with the methodology we consider better.

Stage 2 is complicated: the old methodology will always have (at least cheap) arguments that it is superior to the new. But this doesn't prevent arguments in the other direction. We need to weigh the arguments on each side. We have intuitive ways of doing this, but a formal theory of how we deal with inconsistent information would be nice.

5. The metaphysics of normativity.

Let's get back to the evaluativist picture sketched in Section 3. It has some connection to a reflective-equilibrium picture, on one construal of that.

Not on a construal that says that being justified *consists in* being in reflective equilibrium. (That construal makes the reflective equilibrium view pretty much the same as the coherence theory of justification.) On that construal, the reflective-equilibrium view is totally implausible, in that there are reflective equilibria that we rightly judge as idiotic.

The connection of evaluativism to reflective equilibrium is just that proper methodology consists of striving for reflective equilibrium, not for providing foundations for our beliefs. (This leaves open whether, were equilibrium achieved [!], there would be value in continuing to look for and evaluate other methods. I'm inclined to think so: that even in an equilibrium position where there is no local pressure to change, it would still be of value to look for other nearby equilibria that might be somewhat better. But this issue is somewhat academic: there are always conflicting pressures whose resolutions we should look for.)

Similarly, the evaluativist picture may have something in common with those who advocate a kind of "methodological conservatism", according to which there is value in continuing to believe what one already believes. (After all, advocates of methodological conservatism usually like the Neurath's boat metaphor.) But again, if this means that we are to explain what it is to be justified in believing that p in terms of factors that prominently include actually believing that p, it does not seem a promising approach: there is little merit to a person who conservatively sticks to the story on which the moon landing was a conspiracy, and employs methods designed to immunize this view from criticism.

One could try to fix up the reflective equilibrium and related accounts by adding other factors that are required to make our beliefs justified, but I think that a better approach is to reorient from a focus on *what makes our beliefs justified*. The main problem with the view that reflective equilibrium is "what makes a person's beliefs justified" is that such an approach leaves no room for the evaluator's perspective.

One attempt to get around this is straightforwardly subjectivist: it treats an evaluator's claims about justifiability or reasonableness as claims about the evaluator's epistemic norms. Alternatively, as claims about the norms of the evaluator's community — a kind of "group subjectivism". But whether in individual or group form, this strikes me as not the way to go: it leads to the idea that evaluators with different norms, or from communities with different norms, are just talking at cross purposes when they apparently disagree (or even when they apparently agree). A better way to go is to follow the example of expressivists about other evaluative discourse, e.g. morality.

The term 'expressivism' has been used for a bewildering variety of views, from the non-cognitivism of figures like Ayer, who likened evaluations to cheering and booing, to recent "quasi-realism" whose advocates try to sound so much like normative realists that it's hard to see what the distinctively expressivist feature of their view comes to. Perhaps at this point the label does more harm than good; perhaps 'evaluativist' is a less misleading term, even in the moral case.

And 'expressivism' might be even more misleading in the epistemological case. For a moral expressivist is likely to think that there is a close connection between norms of moral goodness and desires or preferences, and in conversation I've heard it assumed that such a connection to desires or preferences is built into expressivism. This strikes me as inadequate even to the moral case, once one gets beyond 'good': I take the expressivist idea for moral obligation to be that the norms of obligation (or lack of obligation) function more like commands (or permissions) than like desires or preferences. But in the epistemological case there seems even less connection to desires or preferences,[18]

18. There might be some connection at the "second-order" level: in evaluating first-order norms, we may bring in preferences about the weighting of truth-oriented properties. (E.g.: How much risk of falsity balances the chance for truth about a given sort of question? When does a higher chance of approximate truth outweigh a lower chance of exact truth? When does a higher

so anyone who reads any strong such connection into 'expressivism' is bound to find the idea of an "expressivist" epistemology abhorrent. For that reason I'll stick with the term 'evaluativism'.

Here's a thumbnail sketch of the metaphysical aspect of evaluativism as I'll understand it: The key idea is that judgments about what is justified, reasonable, and the like can be divided into two components. One component is a norm of evaluation; the other is a belief in a narrow sense ("pure belief") about what is justified *according to that norm*. Because of the evaluative component, it is natural to declare a normative claim such as "It is reasonable to believe in quarks" *not straightforwardly factual*: in contrast to the straightforwardly factual "There are quarks", the claim about what is reasonable to believe involves epistemological values.

Normative claims have a special kind of perspectival feature that non-normative claims don't share — the perspective being the evaluator's norms. Somewhat similarly, tensed claims have a perspectival feature that untensed ones about 4-dimensional reality don't share, and modal claims have a perspectival feature that non-modal claims about the hyper-universe of possible worlds don't share. (Obviously the normative case is also importantly different from these, in a way to be discussed below.) But just as one can illuminate *tensed* claims by giving an *untensed* account in 4-dimensional terms, and illuminate *modal* claims by giving a *non-modal* account in terms of possible worlds, so one can illuminate *norm-sensitive* claims by an account in language that is *not norm-sensitive*. Gibbard has done so, in terms of his framework of "norm-world pairs". (See Gibbard 1986 and 1990; his later relabeling of "norms" as "hyperplans", in Gibbard 2003, doesn't affect the framework.[19] MacFarlane 2005 and 2014 has offered what can be interpreted as a very similar framework [see note 4], though he focuses less on normativity in applying it.) Norms and worlds are not on par: a norm (hyperplan) is something that assigns to each pair of a belief (or action) and world an evaluation of the belief (or action) at the world (whether just as positive, negative or neutral, or something more fine-grained). Whereas one world is metaphysically privileged (it represents reality), there is no obvious reason to think of one norm (or hyper-plan) as metaphysically privileged. Indeed, it is metaphysically privileged only if the worlds contain "normative facts" that make the norms "correct", and presumably the Gibbard idea was that there is no need for that.

I've stressed that the framework of norm-world pairs is an attempt to model normative language in non-normative language. More specifically, the point of the Gibbard framework is to capture the logic of normativity, as an evaluativist/expressivist sees it: in particular, to show that despite Frege/Geach, the expressivist has nothing to fear about the logic of embedded normative claims. Obviously there are limits on the model: there is no hope of any model accurately capturing the meaning of normative language in non-normative terms, because the normativity itself is a crucial part of their meaning. But attempts to model vague language using non-vague language (e.g. supervaluational semantics or Łukasiewicz semantics) can be illuminating without providing anything like a translation; similarly for modeling tensed or modal language in non-tensed and non-modal terms. And so, I think, in the case of norm-sensitive language.

While the Gibbard/MacFarlane model is valuable, there are specific ways in which one could be misled by it — that is, by the correspondence between norm-sensitive claims (e.g. "You should believe that p" or "'You should believe that p' is true") and claims that explicitly refer to norms in a norm-insensitive way (e.g. "According to norm n, you should believe p" or "'You should believe that p' is true in norm-world pair <n,@>" where @ is the actual world). It is a crucial part of the evaluativist view that a norm-sensitive claim is very different from the corresponding norm-insensitive one. Only to a very limited extent is that difference clarified by the Gibbard framework by itself (as Gibbard himself of course recognizes: he supplements it with illuminating

chance of getting to the truth *quickly* outweigh an overall lower chance of getting to it eventually?)

19. The norms in question are "complete" or maximally detailed norms, just as the worlds are complete or maximally detailed propositions.

remarks on the pragmatics). The framework does allow us to hold *that* there is a difference: it allows us to hold that claims explicitly relativized to norms can have their truth value determined wholly by the world component of a norm-world pair, so that the explicit relativization throws away the norm-sensitivity. But it doesn't tell us what the norm-sensitivity amounts to.

Indeed, the framework itself doesn't say anything about the deep differences there are between norm-sensitivity and the kind of perspectivity one has in the temporal or modal cases. The differences arise from the fact that we can evaluate a temporally sensitive assertion as "objectively correct" if it is true relative to the time intended by the utterer (typically, the time of utterance), and a world-sensitive assertion as "objectively correct" for a possible utterer if it is true at the intended world (typically, the possible utterer's own). Normative discourse doesn't work like that: indeed, for an evaluativist there just is no such thing as objective correctness; there is only correctness in the sense of disquotational truth (where "'p' is true" inherits whatever norm-sensitivity there is in 'p').

The difference between the temporal/modal cases and the normative case becomes especially vivid when one thinks about disagreement. If it were possible for people in different eras to communicate with each other, there wouldn't be disagreement between them when one asserted "The world's human population is now over 3 billion" and the other asserted its negation; similarly if "residents of different possible worlds" could communicate with each other. That's because their "objective correctness" conditions are compatible. But it seems to be of the essence of norm-sensitivity that disagreement doesn't work this way; from an evaluativist perspective, there are no "objective correctness" conditions in this case. There isn't anything within the Gibbard framework of norm-world pairs that *explains* the special pragmatics of disagreement (as, again, Gibbard recognizes), but there is also nothing that rules out there being such a special pragmatics: the framework is just silent on the matter of what constitutes agreement and disagreement. The role of the Gibbard model, as I said, is to get

the logic right, including in particular how the embedding of norm-sensitive claims inside logical operators works.[20]

In sum, the evaluativist framework has it that normative claims have a norm-sensitivity, with a special pragmatic role. This framework allows the following:

- Despite their not being straightforwardly factual, normative claims play an important cognitive role: they certainly aren't merely cheers or boos.

- And they can be rationally evaluated, in part by bringing into consideration other normative judgments and in part

20. These remarks add something to the discussion in Field 2009, which also stressed the pragmatics of disagreement, and presented the view as in the spirit of the "assessor relativism" of MacFarlane 2005, though with a couple of significant differences. That paper offered a motivation for calling evaluativism "relativistic", which is that *the Gibbard modeling* is clearly relativistic: to say that S is true at norm-world pair <n,w> is a notational variant of saying that S is true at world w, relative to norm n. Focusing on the actual world, which is metaphysically fixed, this becomes just truth relative to n.

Quite properly, I refrained from concluding from this that "true relative to n" is the only notion of truth, or even the primary one, in the normative domain. Rather, I took the primary notion of truth to be the disquotational one, where True('p') is equivalent to p, so that "True('p')" inherits whatever norm-sensitivity there is in "p". This means that in a Gibbard model, "True('p')" like p will need to be evaluated at norm-world pairs when 'p' is normative.

However, I did suggest that the Gibbard model suggests a kind of "relativity", not primarily in truth but in ground-level normative notions, but inherited into truth claims from that ground-level relativity. Because of the pragmatic features of normative discourse mentioned above, the relativity had to be of an unusual kind, closer to the "assessor relativism" of MacFarlane 2005 than to prototypical relativism. But the talk of "relativism" here seems optional: one could argue that any "relativism" here is simply an artifact of modeling norm-sensitivity in norm-insensitive language, so that evaluativism itself shouldn't be regarded as a relativist doctrine. I suspect that the issue of relativism isn't clear enough for there to be any point to insisting either that evaluativism involves relativity or that it doesn't. (Either side can accommodate the point that it is sometimes useful to back off the normativity by saying, "Well, it's justified relative these norms, and these norms have such and such advantages": that isn't decisive that "there was some kind of relativity in the normative notions all along".) "Norm-sensitivity" now strikes me as less contentious than "norm-relativity".

by bringing into consideration straightforwardly factual claims.

- We can perfectly sensibly apply the words 'true' and 'false' to normative claims: "If what he said is true, then I shouldn't do X" is perfectly sensible, even if normative claims are among the things he said that are central to my conditional conclusion that I shouldn't do X. (Not only is it sensible to so apply the notions of truth and falsity to normative claims, but disallowing such applications would defeat the main purposes that the notions of truth and falsity serve.)
- Connected with the last point, it's natural to say that I *believe* such normative claims: it's just that this isn't *pure* belief; it has in addition an evaluative element.
- Moreover, many normative claims clearly have what we might call *counterfactual objectivity*: we can properly say that I wouldn't be justified in believing the Earth flat even if I had very different epistemic standards that dictated such belief; indeed, we can point out that it is inconsistent with the standards we accept to positively evaluate belief in a flat earth by people with those standards.

The bulleted claims go some of the way toward "quasi-realism": *much* of what the realist says can be said by the evaluativist/expressivist as well. But not all: in particular, and in contrast to Blackburn 1993 and the more recent Gibbard (e.g. 2003), I think that there is an issue of objectivity that goes beyond the issue of counterfactual objectivity, and that the realist believes in that further sort of objectivity but the expressivist, quite properly, doesn't.[21] It has to do with the fact that, on the expressivist view, our normative claims *arise out of* our norms but don't posit a counterfactual dependence on our acceptance of those norms.

Sharon Street (2011) has complained, quite plausibly, that recent quasi-realist views have gone so far toward accepting what the realist says that they lose any epistemological advantage over realism. (In the present context, the kind of epistemological advantage concerns what I earlier called *meta-skepticism*: skepticism about claims of justification or reasonableness, which, as we've seen, can indirectly lead to ground-level doubts, e.g. about the external world.) But the current proposal is not quasi-realist in that sense: unlike the quasi-realist, the evaluativist makes no claim to be just like the realist as regards objectivity. After all, as far as the metaphysics goes, the evaluativist is in exact agreement with the individual or group subjectivist (in Street's terminology, the constructivist). The only difference is in the way that metaphysics is accommodated in the treatment of language. I regard the evaluativist view of how to accommodate it as far more natural than the subjectivist/constructivist, but on matters of non-counterfactual objectivity and on consequent matters of epistemology they seem to me precisely the same.[22]

I concede that there is more to be said about both the distinction between norms for valuation and "pure beliefs", and the kind of non-counterfactual objectivity that is connected to it.[23] Quasi-realists seem

21. Somewhat related to this: advocates of quasi-realism often say that accepting their doctrine does not in any way affect ground-level normative practice; whereas a main theme of the present paper is that evaluativism does have such effects.

22. A referee has suggested that the epistemological challenge that Street is addressing (the "reliability challenge"; I won't take the space to explain it) is generated solely by the counterfactual aspect of objectivity. I disagree: for an evaluativist, the challenge is met by our *actual* acceptance of norms that apply even in counterfactual circumstances where we accept different norms. (Some of the literature on the reliability challenge engenders confusion over this, by talking about "mind-independence" or "attitude-independence" ambiguously: a claim can be sensitive to our actual attitudes, in the Gibbard-MacFarlane sense, without being counterfactually dependent on them.)

23. The difficulty of achieving complete clarity on the non-counterfactual notion of objectivity has some parallel in the difficulty of achieving complete clarity on the notion of determinacy. Just as the truism that 'Joe is bald' is true if and only if Joe is bald would seem to leave open that there is a question of whether there is a *determinate* fact of the matter whether Joe is bald, similarly the truism that 'the joke was funny' is true if and only if the joke was funny would seem to leave open that there is a question of whether there is an *objective*

to question that there is such a further aspect to objectivity, but as Crispin Wright has often argued (e.g. Wright 1992), this is exceptionally implausible: for instance, claims about what is funny can have counterfactual objectivity, but it's hard to believe that they are objective in any very deep sense. Of course the evaluativist about norms in morality and/or epistemology will grant that there are key differences between the form that the non-objectivity takes in these cases and in the case of humor — for instance, our epistemological and moral norms are way more highly structured than is our sense of humor, and far more deeply entwined with our goals. But my point wasn't to assimilate the normative to the comic, but simply to say that there is far more to objectivity than counterfactual objectivity.

I hope it is also clear that on my view there are deep differences between epistemological normativity and moral normativity. Whereas there is presumably considerable indirect connection between moral norms and preferences about what kind of world one wants to live in, it's hard to see much analogous connection in the epistemological case, at least at the first-order level (see note 18). More fundamentally, in evaluating either epistemological or moral norms in terms of how well they satisfy given factual desiderata, one needs to use epistemological norms but doesn't need to use moral norms. This last difference is what makes reliabilist-like views tempting: it makes tempting that what matters to epistemological goodness is a purely factual matter, something like the "truth-conduciveness" of the rules by which a belief is formed and retained. But succumbing to the temptation isn't inevitable, and I've mentioned some problems with doing so. Above all, I don't think there's much hope in making the idea of "truth-conduciveness" at all clear; and even insofar as it is clear, *exclusive* focus on actual truth-conduciveness demotes too much the "internalist" features of epistemological evaluation.

fact of the matter whether the joke was funny. There may be more than just a parallel here: maybe indeterminacy is just a special kind of non-objectivity. In any case, the positive task of explaining either determinacy or objectivity with complete clarity is difficult.

A final word on objectivity: it's best to view it as coming in degrees. There are some norms of evaluation which have so much going for them that for many purposes they can be regarded as objective: these might include

(i) some logical norms,

(ii) some norms of comparative evidence for statistical hypotheses (e.g. "If an experiment E was performed and led to result R, then $E \wedge R$ favors a statistical hypothesis that gives R higher probability in circumstances E over another statistical hypothesis that gives it lower probability"), and/or

(iii) some form of the "Principal Principle".

I don't mean that such norms are entirely uncontroversial: they aren't, especially when it comes to the detailed formulation. (There are debates about the details of "the correct logic" and about the best form of the Principal Principle; and views of statistical inference that seem in conflict with the comparative likelihood rule (ii) are widespread, though not as widespread as they once were.) I'm doubtful that the controversies over such rules ought to count as objective in the way that controversies over the existence of gravitational waves are,[24] but if someone wants to argue that these are perfectly objective matters, I'm not going to put up a fight. In contrast, the *non-comparative* evaluation of hypotheses (even statistical ones), and the evaluation (even comparative) of hypotheses that are not purely statistical, both make a far more serious use of either prior credences (in a Bayesian framework) or something that plays a similar role to that in a non-Bayesian framework. (The non-comparative evaluation of hypotheses involves something like a prior credence function over the space of alternative hypotheses; and any evaluation of hypotheses that aren't purely

24. Even when the choice between two norms isn't objective, there can be considerable advantages of one norm over the other; rational debate over the norm consists in pointing out such comparative advantages.

statistical involves something like prior credences of auxiliary hypotheses.) It is primarily here that I'd want to insist on a significant level of non-objectivity. Though even here it is a matter of degree (where the degrees are vague and not linear-ordered): if a claim is given the same evaluation by all methods we can take seriously, then it should count as highly objective. My view is that dichotomizing between the "objective" and the "non-objective", in a context-independent way, isn't very useful. (Understood as contextual, a vague dichotomy makes sense: there's nothing wrong with counting claims with a low degree of whatever kind of non-objectivity is salient in the context as "objective", as long as we're clear that that's all we're doing.)

6. Skepticism again.

Let's look a little more at the impact of the evaluativist/expressivist picture on epistemology. I've mentioned skepticism already, but will say a bit more about it in this section. In the final sections I'll give two more examples, the first of which has a fairly close connection to issues discussed in Sections 3 and 4, and the second of which is somewhat further removed.

I noted early on (footnote 1) that the discussion in this paper does not target all forms of skepticism. And in addition to the examples mentioned in that footnote, it does not directly challenge skepticism based on Benacerraf-style arguments against certain forms of Platonism: there the skepticism is based not on lack of initial justification but on an apparent undercutting of that justification. (The discussion in this paper does challenge the analog of Benacerraf-style arguments for normativity, but only by challenging the normative realism on which they are based; it is a separate question what to say about specific forms of mathematical Platonism.)

More important, the normative anti-realism recommended in this paper does not totally foreclose even the kinds of skepticism that were targeted: it merely makes them far less well-motivated. Part of the reason why those forms of skepticism can seem compelling is that even the forms of normative realism designed to blunt it, such as the "entitlements" idea discussed in Section 1, seem to let it in through the back door via skepticism about the entitlements. The normative anti-realism does foreclose *that meta-justificational route* to skepticism. Nonetheless, it would be possible to adhere to the anti-realism while adopting norms that would lead to skepticism by other routes.

One such norm ("Complete Open-Mindedness") is that we should take seriously any hypothesis that is ever suggested, no matter how silly: keep it as a live possibility unless one has non-question-begging reasons to eliminate it. Adopting this methodology would allow for initial justification in claims about the external world; but once the brain-in-vat (or even the Cartesian demon) hypothesis is suggested, the methodology would no longer allow for the acceptance of claims about the physical world. (At least, not unless non-question-begging reasons could be provided to eliminate it, and that presumably is impossible.)

The obvious anti-realist response to this is simply that we don't accept Complete Open-Mindedness, and shouldn't. Descartes' powerful writing persuaded many philosophers to take an application of that norm at least *somewhat* seriously: seriously enough to have a bad conscience about going on as before without providing reasons against the demon. But the norm also has consequences that not even Descartes could have taken seriously. For instance, consider the hypothesis that the world is as scientists believe, and will remain so as long as no one hops around the South Pole 91 times on one leg while singing "Twist and Shout" in falsetto; whereas if someone does that, global warming will be reversed and plenty of food will become available to everyone in perpetuity.[25] I venture to say that even having raised this hypothesis, no one will ever test it, despite the benefits to mankind it promises. That's because we don't accept the norm of Complete Open-Mindedness. And we shouldn't: our norms about what norms to accept are such as to dictate that accepting it would be a thoroughly bad idea.

25. I recall Hilary Putnam giving a similar example in a class many years ago.

Again, a normative realist could raise a meta-justificational issue here: what is the "objective justification" (not based on meta-norms that could be questioned) of the claim that we shouldn't accept the Complete Open-Mindedness norm? An advantage of the evaluationist metaphysics is that it obviates that question.

There may still be an issue for the normative anti-realist: given that we don't accept the Complete Open-Mindedness norm, what norm of comparable generality do we accept? The question has a presupposition: that we do accept norms of comparable generality. I suspect that that presupposition is false, though this is not an issue on which I want to take a stand.[26] I suspect that we can't do much better than say, "We ought to be open-minded, within reason, but each person must decide for him- or herself just what alternative hypotheses are worth taking seriously."

7. Epistemic rules.

In addition to skepticism, there are other ways in which the evaluativist/expressivist picture affects ground-level epistemology. For instance, I think that the evaluativist picture undercuts a certain conception of "rules of rational belief". That conception tends to lead, in particular, to a conception of *fundamental* rules of rational belief *that are immune from rational revision*. But what I see as the basic error in the conception comes before that. The picture I reject is stated very clearly in the opening paragraph of Paul Boghossian 2008 (and though the burden of that paper is that the conception of rules described there seems to lead to paradox, I don't think he gets at what seems to me to be wrong in the conception). He says:

> … we have to try to figure out what is true from the evidence available to us. To do this, we rely on a set of epistemic *rules* that tell us in some general way what it would be most *rational* to believe under various epistemic circumstances. We *reason* about what to believe; and we do so by relying on a set of rules. [Boghossian 2008, p. 472]

So the rules here are rules governing what is *rational to believe*; but they also serve a somewhat more descriptive function, in that they are *involved in the reasoning processes* of someone who rationally believes. I don't think that these need be the same; indeed, I think it easier to motivate talk of rules involved in the reasoning processes of believers (rational and otherwise) than to motivate talk of rules of what it's rational to believe.

I'm sympathetic to Boghossian's claim that we reason by relying on a set of rules. Talk of rules is probably important in the quasi-descriptive phase of epistemology: that is, in the task of giving idealized descriptions of how, at a given time, an actual or possible agent (or maybe an actual or possible community of agents) would function in absence of "performance errors". (By "rules" I don't mean just explicitly formulated rules that the agent consults; like Boghossian, I mean to include rules that the agent follows blindly.) As I've said, there needn't be a clear model-independent distinction between what counts as a performance error and what doesn't. We model epistemological behavior by idealized descriptions, where behavior that doesn't fit the description is counted as performance error; what counts as a performance error on one idealization may be built into the idealization on another, and different such "levels of idealization" may be useful for different purposes. The rules are level-dependent.

It may well be that in any idealized descriptive model, some rules will be fundamental in the sense that the model doesn't allow any

26. The connection between evaluativism and the issue of whether to look for completely general norms is a weak one. I do think there is more pressure to aspire to complete generality in norms if one takes there to be a metaphysically based correctness of norms than if one simply takes our norms to be products of our biological and cultural adaptation to our circumstances, for which talk of correctness doesn't arise. But I acknowledge that there are particularist realists who very much downplay the pressures toward systematicity from a realist perspective, and also that achieving complete generality might have some kind of appeal from the anti-realist perspective as well as from the realist.

considerations to undermine them. (Such fundamental rules, if they exist, needn't be deterministic: they could allow that in certain circumstances an "internal coin flip" dictates how to proceed.) Allowing for fundamental rules in this sense might seem to make a problem for the idea, implicit in Sections 3 and 4, that even basic deductive and inductive and perceptual methodology can be challenged under certain circumstances: don't we want to have epistemological models that build in logical and inductive and perceptual rules, but also models that deal with the process of challenging them? One way to deal with this is to simultaneously employ multiple models, at different levels of idealization. Consider, as a crude illustration, a two-tape Turing machine where the top tape contains instructions for inductive method M. The method on the top tape is appealed to constantly in rewriting the bottom tape used for ordinary practice. Only in very exceptional circumstances do the overall rules of operation of the machine (the machine table, which I'm imagining to be built into the architecture and thus not explicitly appealed to) dictate rewriting the top tape. From the point of view of the overall Turing machine rules, M is merely a default program that has defeaters. But anyone who wants to give a manageable description of the behavior of the machine at a given time will appeal not to the machine table but to the default rules in M.

To repeat, I'm conceding that a quasi-descriptive epistemology will posit rules, even perhaps rules that *on that level of description* are fundamental. And, if the quasi-description is of an epistemologically good agent, perhaps we could call the rules it posits (at that level of description) "rules of rational belief" (at that level of description). But speaking this way is potentially misleading, on a number of levels.

For one thing, it might suggest that the notion of rationality will appear in the rules; and indeed, Boghossian thinks it does. For instance, he formulates the rule of Modus Ponens as "If you are rationally permitted to believe both that p and that 'if p then q', then you are *prima facie* rationally permitted to believe that q" (Boghossian 2008, p. 472). But on a picture that separates the quasi-descriptive from the evaluative, the rules will not themselves employ normative concepts.

For another thing, speaking of "rules of rational belief" strongly suggests that the *only* factor in an evaluation of the rationality of a given belief is whether it was produced via a good rule and without performance error. That is unobvious.

Another problem with the "rules of rational belief" terminology is that it strongly suggests that there is a set of optimal rules: good agents are the ones who follow those rules, or some approximation to them. If there is such a set of optimal rules, which I doubt, that needs an argument, and nothing in the quasi-descriptive picture suggests any such argument.

Even more important than the optimality issue is the suggestion that there are rules of rational belief *that are fundamental in a model-independent sense*. Nothing about the role of rules in the quasi-descriptive picture of an epistemologically good agent provides the basis for such a claim. When there are multiple levels of description in terms of rules, how are we to decide which level corresponds to "the rules of rational belief"? Indeed, in the toy model of the 2-tape Turing machine, it isn't at all clear what should count as a "rule of rationality".

That is especially so if one makes the assumption (which isn't explicit in the Boghossian quote, but often taken as part of the "rules of rational belief" picture) that the "rules of rationality" are *rationally indefeasible*. If we take the rules of rationality to include typical logical or inductive methods, such as might be included in the instructions in the top tape at a given time, then the toy model is one in which the rules of rationality can change. And they might change in an intuitively rational way: the change might come about by normal operation of the machine, rather than by a malfunction, and this normal operation might be an intuitively good one in that

(i) it changes the method on the top tape only in situations where a deficiency in that method has been exposed, and

(ii) the kind of changes it makes seem well-designed for correcting those deficiencies.

So, in the toy model, the rules of M aren't a candidate for indefeasible rational rules. And the only other rules that played a role in that description were the rules (not explicitly represented ones) that are built into the machine table. But the machine table is far too "low-level" to be naturally viewed as describing "rules of rationality".

Perhaps it will be said that if the machine table operates well in its revision of the methods on the top tape, it will accord with certain heuristics (that it needn't explicitly represent). But such heuristics needn't include general deductive and inductive methods; they can be merely rules for changing the top tape; they aren't a substitute for what's on the top tape, so it would be odd to confine the term 'rule of rationality' to them. (Moreover, evaluating a change in M as rational in a given case needn't require a judgment that the method used to change it would in general lead to rational changes.)

The moral then is that we need to cleanly separate the quasi-descriptive task from the evaluative (while granting that the evaluation of the methods produced in quasi-descriptive accounts has considerable bearing on the evaluation of beliefs). Once we make this separation, certain arguments for the rational indefeasibility of methods we take to be good ones evaporate.[27]

8. Rational constraints.

Many philosophers who have gone some way toward the sort of evaluativism I've been recommending don't seem to me to have broken from the grip of the idea that epistemology is engaged in "describing the normative facts".

For example, many Bayesians who reject the idea that for every epistemological situation there is a uniquely rational epistemic credence function to have in that situation think that it's important to ask what the "constraints" are on a credence function — which credence functions accord with "the rules of rationality" and which don't. From an evaluativist perspective, it's hard to find anything sensible that this can mean. It is of course true that there are credence functions that it would be idiotic to employ: e.g. the aforementioned credence function that protects the view that the moon landing was a hoax from all counter-evidence. But that doesn't mean that there's any point in looking for a "rule of rationality" that "constrains" us not to employ such a credence function (in some metaphysical sense of constraining that is difficult to make sense of, except perhaps in a theological sense of divine punishment). As Dick Jeffrey once said, "The fact that it is legal to wear chain mail on city buses has not filled them with clanking multitudes" (Jeffrey 1983, p. 145).

One might, I suppose, try to interpret talk of "rational constraints" more subjectively, in terms of the features that we would require of a credence function if we are to *deem* it rational (or in terms of the features that an evaluator *that we would deem good* would require of a credence function if the evaluator is to deem it rational). Three points about this: First, our evaluation of credence functions isn't the yes/no affair that talk of satisfying or failing to satisfy rational constraints would suggest: for instance, a credence function might be pretty good, but slightly deficient in not allowing enough credence to certain kinds of hypothesis. Second, even if constraint talk were weakened to accommodate this (e.g. allowing "constraint" to come in degrees), the suggestion seems to require that our evaluations are or ought to be systematizable in a very particular way: that we employ general constraints-to-a-certain-degree on credence functions, and evaluate a credence function as rational to precisely the extent that it meets these general constraints. As I said near the end of Section 6, this assumption

27. There are related salutary effects too. For instance, the argument that Boghossian has given in several papers for regarding standard deductive rules like Modus Ponens as "basic rules of rationality" seems to be based on the supposed inevitability of the fundamental employment of such rules in a descriptive account of any agent that reasons properly. I'm skeptical that it is inevitable, at the basic level: a Turing machine can be programmed to infer according to Modus Ponens, but its basic rules don't include Modus Ponens (though of course *we* use Modus Ponens in reasoning about what such a machine would do). As noted, we presumably don't want to call the basic Turing machine rules (or analogously, the rules governing the evolution of the state of a neural network in humans) "rules of rationality", but the point is more general: maybe the "basic rules of rationality", if there are such, are not rules that involve logical notions at all.

seems far from obvious. Third and most important, I doubt that the subjectivist reading captures the spirit behind typical talk of rational constraints: after all, most of us recognize that our own judgments as to what is rational aren't the last word, so that a credence function that doesn't meet the "constraint" of what we *deem* rational might turn out to have real advantages.[28]

References

Blackburn, Simon 1993. *Essays in Quasi-Realism*. Oxford University Press.

Boghossian, Paul 2003. "Blind Reasoning". *Proceedings of the Aristotelian Society Supplement* 77: 225–48.

------------------- 2008. "Epistemic Rules". *The Journal of Philosophy* 105: 472–500.

Brown, Bryson, and Graham Priest 2004. "Chunk and Permeate, a Paraconsistent Inference Strategy. Part I: The Infinitesimal Calculus". *Journal of Philosophical Logic* 33: 379–88.

Carroll, Lewis 1895. "What the Tortoise Said to Achilles". *Mind* 4: 278–80.

Enoch, David, and Joshua Schechter 2006. "Meaning and Justification: The Case of Modus Ponens". *Noûs* 40: 687–715.

Feyerabend, Paul 1975. *Against Method*. New Left Books.

Field, Hartry 2000. "Apriority as an Evaluative Notion". In Paul Boghossian and Christopher Peacocke, eds., *New Essays on the A Priori* (Oxford University Press), pp. 117–49.

--------------- 2009. "Epistemology Without Metaphysics". *Philosophical Studies* 143: 249–90.

--------------- 2015. "What Is Logical Validity?". In Colin R. Caret and Ole T. Hjortland, eds., *Foundations of Logical Consequence* (Oxford University Press), pp. 33–70.

28. I've received a great deal of helpful commentary on previous drafts, which have led to big improvements. Thanks to Paul Boghossian, David Enoch, Jim Pryor, Stephen Schiffer, Elliott Sober, Lisa Warenski, Crispin Wright, and two anonymous referees.

Gibbard, Allan 1986. "An Expressivistic Theory of Normative Discourse". *Ethics* 96: 472–85.

------------------ 1990. *Wise Choices, Apt Feelings: A Theory of Normative Judgment*. Harvard University Press.

------------------ 2003. *Thinking How to Live*. Harvard University Press.

Jeffrey, Richard 1983. "Bayesianism with a Human Face". In John Earman, ed., *Testing Scientific Theories* (University of Minnesota Press), pp. 133–56.

Kolodny, Niko, and John MacFarlane 2010. "Ifs and Oughts". *The Journal of Philosophy* 107: 115–43.

Kratzer, Angelika 2012. *Modals and Conditionals*. Oxford University Press.

Lewis, David 1971. "Immodest Inductive Methods". *Philosophy of Science* 38: 54–63.

---------------- 1974. "Spielman and Lewis on Inductive Immodesty". *Philosophy of Science* 41: 84–5.

MacFarlane, John 2005. "Making Sense of Relative Truth". *Proceedings of the Aristotelian Society* 105: 321–39.

--------------------- 2014. *Assessment Sensitivity: Relative Truth and Its Applications*. Oxford University Press.

McGee, Vann 1985. "A Counterexample to Modus Ponens". *The Journal of Philosophy* 82: 462–71.

Pollock, John 1987. "Epistemic Norms". *Synthese* 71: 61–95.

Reichenbach, Hans 1938. *Experience and Prediction: An Analysis of the Foundations and the Structure of Knowledge*. University of Chicago Press.

Salmon, Wesley C. 1967. *Foundations of Scientific Inference*. University of Pittsburgh Press.

Street, Sharon 2011. "Mind-Independence Without the Mystery: Why Quasi-Realists Can't Have It Both Ways". *Oxford Studies in Meta-Ethics* 6: 1–32.

Williamson, Timothy 2007. *The Philosophy of Philosophy*. Blackwell Publishing.

Wright, Crispin 1992. *Truth and Objectivity*. Harvard University Press.

-------------------- 2014. "On Epistemic Entitlement (II): Welfare State Epistemology". In Dylan Dodd and Elia Zardini, eds., *Scepticism and Perceptual Justification* (Oxford University Press), pp. 213–47.