

To appear in *Philosophical Studies*, 2009. Online publication January 2009, available on [springerlink.com: http://dx.doi.org/10.1007/s11098-009-9338-1](http://dx.doi.org/10.1007/s11098-009-9338-1).

## **Epistemology without Metaphysics**

Hartry Field

A common picture of justification among epistemologists is that typically when a person is looking at something red, her sense impressions pump in a certain amount of justification for the belief that there is something red in front of her; but that there can be contrary considerations (e.g. testimony by others that there is nothing red there, at least when backed by evidence that the testimony is reliable) that may pump some of this justification out. In addition, the justification provided by the senses can be fully or partially undercut, say by evidence that the lighting may be bad: this involves creating a leak (perhaps only a small one) in the pipe from sense impressions to belief, so that not all of the justification gets through. On this picture, the job of the epistemologist is to come up with an epistemological dipstick that will measure what overall level of justification we end up with in any given situation. (Presumably the “fluid” to be measured is immaterial, so it takes advanced training in recent epistemological techniques to come up with an accurate dipstick.)

Of course there are plenty of variations in the details of this picture. For instance, it may be debated what exactly are the sources of the justificatory fluid. (Does testimony unaided by evidence of its reliability produce justification? Do “logical intuitions” produce it? And so on.) Indeed, coherentists claim that the idea of sources has to be broadened: build a complex enough array of pipes and the fluid will automatically appear to fill them. There are also debates about the fluid dynamics. For instance, the question of what exactly are the circumstances under which a valid argument “transmits justification” from premises to conclusion has been much discussed in recent years.

The aim of this paper is to offer an alternative picture. It is one I’ve tried to set out before, but I don’t think I’ve conveyed it very successfully, and hope I will be forgiven another attempt.<sup>1</sup>

---

<sup>1</sup> My first published attempt in this direction (aside from very sketchy remarks near the end of Field 1982) was a discussion of nonfactualism about normativity generally, in Sections 3 and 4 of Field 1994 (republished with significant corrections in Field 2001), which was much influenced by Gibbard 1990. The application to epistemology was discussed in Field 1998b, 2000 and 2005. The view is at least somewhat similar in spirit to views recently expressed by

**1. Expressivist relativism.** The basic idea I want to develop is that regarding a belief as justified (or reasonable, or rational, or whatever)<sup>2</sup> is evaluating it, and evaluations aren't straightforwardly factual. More fully: (i) regarding a belief as *epistemically* justified or *epistemically* reasonable is evaluating it *from an epistemological perspective*, that is, from a perspective that rules out factors we deem outside the scope of epistemology (such as the prudential considerations that might “non-epistemically justify” someone believing that his partner is faithful); and (ii) epistemic evaluations, like other evaluations, aren't straightforwardly factual.

The analogous position about moral evaluations is familiar, though there are controversies about how best to develop it. I think that most of the reasons for the position in the moral case extend to the epistemological case as well.<sup>3</sup> There are additional difficulties to be overcome in the epistemological case; these will be mentioned in Section 3 and further discussed in Section 11. The final section will be a brief attempt to convey some of the benefits of the position to epistemological practice. But much of the paper will be concerned with elaborating the general idea that evaluations (whether epistemological, moral, aesthetic, or any other) have a

---

Max Kölbel 2004 and John MacFarlane 2005, and I have benefited from thinking about these works, as well as from discussions with numerous students and colleagues including Sinan Dogramaci, Gary Ebbs, Andy Egan, David Enoch, Melis Erdur, Dana Evan, Paul Horwich, Matt Kotzen, Maria Lasonen-Aarnio, Jim Pryor, Karl Schafer, Stephen Schiffer, Ted Sider, Sharon Street, David Velleman, Lisa Warenski, Tim Williamson, and Crispin Wright. Paul Boghossian and Stewart Cohen deserve special mention: Boghossian's persistent skepticism about my views has led to a considerable sharpening of them, and Cohen's extensive correspondence about the near-final draft led to some substantial improvements. Andy Egan and Seth Yalcin have pointed out connections between some of the ideas in the paper and recent work on epistemic modals, by them and others, but I've decided not to expand an already long paper to pursue this.

<sup>2</sup> While distinctions between these terms are possible, I don't believe there is any standard such distinction. Perhaps the main difference is that 'justified' seems ambiguous: calling a belief justified often means that it's reasonable, but occasionally means that there is something which serves as a justification for it; illicitly sliding between these two readings is a common source of epistemological chicanery. For this and other reasons, I prefer 'reasonable', and will generally use it in the rest of the paper.

<sup>3</sup> For instance, (i) the usual metaphysical (Humean) worry, that there seems no room for “straightforward normative facts” on a naturalistic world-view; (ii) the associated epistemological (Benacerraf-style) worry that access to them is impossible (which is compounded by the fact that there is substantially greater disagreement about normative matters than about mathematical); (iii) the worry that the relation to norms is not only non-naturalistic, but “queer” in the sense that it's supposed to somehow motivate one to reason in a certain way. (Another worry related to (iii) will be mentioned near the end of Section 8.)

kind of not-straightforwardly-factual status.

That idea immediately raises at least three problems, which are related:

- (I) What does it mean to say that these evaluations aren't straightforwardly factual?
- (II) How can the view be extended to embedded constructions involving 'justifies' or 'reasonable', such as the claim (or judgement) that *if it's reasonable to believe that the Earth is round then it's reasonable to believe that people on the other side will fall off?*
- (III) How can the view accommodate the obvious fact that people can debate claims about what's reasonable (including claims in which 'reasonable' occurs highly embedded), and how can it allow "straightforwardly factual" claims (for instance, about physics or about human fallibility) to have a role in such debates?

Problems analogous to these were very serious worries for many earlier views about evaluative discourse that have similar motivations to mine.

The answer to these questions is going to involve *a kind of* relativism: in some sense, evaluative claims involve a free parameter, for a norm of assessment. But there is a big difference between this kind of free parameter and the free parameters involved in ordinary context-relative sentences like 'It is raining'. (Actually there are two big differences, but I'll defer the second to the next section.) The difference is that with 'It is raining', the speaker intends to make a claim about a specific spatio-temporal location (or if you like, a claim whose correctness is determined by what's happening at a specific location that the speaker intends).<sup>4</sup> For instance, the speaker may intend to say that it is raining *where he is now*, or that it is raining *at some other location that he intends or assumes is contextually determined*; in some cases it will be indeterminate which location he had in mind, but then no determinate thought is expressed. In an evaluative claim, on the other hand, one doesn't intend to be making a claim

---

<sup>4</sup> I take no stand on whether the location is needed to determine which proposition is expressed, or is instead used as a parameter in terms of which the proposition is evaluated. This distinction (between "indexicality" and "non-indexical contextualism", in current terminology) strikes me as more a matter of legislation on the use of 'proposition' than a matter of substance.

about a specific norm (or a claim that is to be evaluated for correctness by looking at a specific norm that the speaker intends): a claim about what is *justified according to a specific norm* would be straightforwardly factual, with no evaluative force. (It would encourage the Moore-like response “Sure it’s justified *according to that norm*; but is it *justified?*”)<sup>5</sup> What I’m advocating for normative terms is very different from contextual relativism, *so* different that in my 1994 paper I decided not to call it ‘relativism’ at all, and to label it a kind of expressivism (though one very different from old-fashioned versions of expressivism, in that it gives evaluative statements a cognitive role). But John MacFarlane (2005) has recently introduced the term ‘assessor-relativism’ for what seems at first blush to be just this sort of thing. My understanding may not fully coincide with his—more on this below, especially in Section 10—but I will use his term nonetheless.

The term ‘expressivism’ seems appropriate too, since the view has it that an evaluative sentence expresses a mental state that is a resultant of norms and factual beliefs. In fact, the view seems to be a notational variant of Allan Gibbard’s view (1990, Chapter 5) that evaluative claims express propositions in an extended sense: not just sets of possible worlds, but sets of norm-world pairs. If an evaluative claim *A* expresses an extended proposition consisting of a set of norm-world pairs  $\langle n, w \rangle$ , then *A* is something that can be true at a world *w* relative to a norm *n*. In a typical assertion we are making a claim about the actual world @—we are “filling the world slot with @”. If Gibbard intended there to be a metaphysically privileged norm playing a role analogous to that of the actual world, it would be a metaphysically realist view of the sort he clearly was trying to provide an alternative to; so the only sensible interpretation of Gibbard is that he intended to leave the norm parameter unfixed. The same normative proposition can be true at the actual world relative to some norms but not others, so that different speakers who agree on all the relevant facts can still evaluate it in different ways by employing different norms in making their evaluations. And none of their conflicting evaluations would be metaphysically privileged over the others. That’s assessor relativism as I’ll understand it.

So the view has both elements associated with relativism (though not of the contextual

---

<sup>5</sup> Indeed, there typically isn’t an intended norm, or even a small set of norms that are candidates for being intended: one’s normative views tend to be far looser than would be required for this.

sort) and elements associated with expressivism (though not of the non-cognitivist variety in which “no proposition is expressed” or in which one’s factual beliefs don’t enter into normative evaluation). I’m tempted by the term ‘expressivist relativism’ (or maybe ‘relativist expressivism’); but ‘evaluativism’ is shorter, so I’ll use that.<sup>6</sup>

**2. Description or revision?** Will evaluativism be proposed as a description of the ordinary meaning of sentences about justification, or as a recommendation for the revision of ordinary practice? The answer is that I take no stand on what the ordinary speaker means—indeed, I doubt that that’s a clear question. So I certainly don’t claim to be describing ordinary practice or meaning; but I also don’t claim to be offering a revision of ordinary practice or meaning, since that would require supposing that ordinary practice or meaning differs from what I’m suggesting. What I do claim to be revisionary about is the philosophical views of normative realists (in particular, epistemological realists). These people believe in a kind of “metaphysical justification”. (Usually they attribute this belief to the ordinary speaker, as “part of what the ordinary speaker means” by ‘justification’, though that sociological claim is in principle separable from their view.) I think that to the extent that sense can be given to this notion of metaphysical justification, there’s no such thing. So my view is that *to the extent that* ordinary people are committed to metaphysical justification, they are in error; and to the extent that “the meanings of their words is so committed”, these meanings are founded on error and must be

---

<sup>6</sup> The focus of the paper is on justification or reasonableness rather than knowledge, but I’m inclined toward some kind of evaluativism for knowledge too. One approach to knowledge takes it as involving justification to a sufficiently high degree that is undefeated. It seems to be widely accepted today that *the required degree* is somehow relative—usually this is taken as some sort of contextual relativity, but see Chrisman 2007 for a more expressivist variant similar in spirit to the kind of expressivism advocated in this paper. Most of this recent literature however involves relativism or expressivism only about the degree of justification required for knowledge; whereas I’m applying it to the standards of justification itself.

Of course, it’s controversial that knowledge requires justification at all; and if it doesn’t, the case of knowledge would require a quite separate discussion. But whichever way one goes on that, it seems to me that knowledge is a normative concept, and that straight naturalistic accounts (e.g. in terms of reliability) simply don’t do justice to it; and I think that something in the broad ballpark of what I say here would still apply. But that’s too big a matter to discuss here.

replaced. At the same time, I want to employ a notion that I think is natural to call ‘justification’, which I think shares much of the role that that term plays in ordinary practice but does not have any such metaphysical commitments. The task as I see it is to lay out how such a notion might work, and to try to argue that most of ordinary practice (though not all the practice of normative realist philosophers) could survive using just such an ametaphysical notion of justification. To the extent that this is successful, its success could be used to make *some* kind of case for the view that the best interpretation of ordinary speakers is that they are not committed to anything more than the ametaphysical notion; but, as I said, I don’t want to take a stand one way or the other on how to interpret the ordinary speaker. (Indeed, I don’t regard questions about “what the ordinary speaker means” as clear, though this latter is based on controversial views in philosophy of language that I don’t want to get into here.)

The proposal, then, adopts the metaphysics of the error theorist, without commitment one way or the other to the error theorist’s claim to be a revolutionary. But if you follow the error theory more fully, by holding that the ordinary speaker is in error or that the meanings of ordinary evaluative words is based on error, you can take the proposal as one for how to talk once we have recognized the error.

My view is that adopting the metaphysics of the error theorists needn’t result in doing much violence to ordinary discourse that involves normative terms: ordinary normative discourse can, to a very large extent, be accommodated. This has *some of* the flavor of what Simon Blackburn (1993) has called quasi-realism. But unlike quasi-realism, this view does not attempt to mimic the normative realist (even in “ground level normative” as opposed to meta-normative discourse).<sup>7</sup> The view is that we can mimic a great deal of ordinary practice *without* mimicking

---

<sup>7</sup> Like Dworkin (1996) I am skeptical that a hard and fast line can be drawn between ground-level normative claims (“internal claims”) and meta-normative claims (“external claims”): I’m inclined to agree that any sort of anti-realism about normativity is bound to affect ground level practice. It needn’t affect it by actually *denying* the realist’s claims: if a realist wants to say that it is a straightforward objective fact, out there, independent of us, that one shouldn’t kick dogs for the fun of it (cf. Gibbard 2003, p. 186), the appropriate reaction of the anti-realist isn’t to take a stand on the claim but to question the rhetoric and give a story on which the rhetoric is at the very least highly one-sided and in most contexts thoroughly misleading. (In the cited passage Gibbard advocates an *internal adequacy thesis* that seems to accept the hard and fast line, and the quasi-realist program. However, he takes the judgement about dogs to be ambiguous between an internal and external reading, and says that the rhetoric is “sumptuous” on the internal, so I’m not sure the extent of real disagreement. Still, adopting the aim of sounding like the realist seems to me to distort the enterprise.)

normative realism.

I remarked earlier that there are two big differences between the kind of relativity involved in claims about justification and the kind of relativity involved in ordinary contextual sentences like ‘It is raining’. The first, which I’ve mentioned but haven’t yet really explained, is contextual relativity *versus* assessor relativity. The second, implicit in the last few paragraphs, is that while the relativity in ‘It is raining’ is entirely uncontroversial, the relativity in epistemological judgements is far from that: it is certainly contested by epistemological realists, and at least arguably goes against the opinions of the person on the street.

This difference is independent of the difference about contextual relativity *versus* assessor relativity. Indeed, 20<sup>th</sup> century physics is full of examples of controversial (or anyway, once-controversial, and still not universally known) relativity that seem contextual. The example most commonly discussed in this connection (Field 1994, Harman 1996, Boghossian 2006b) is the relativity in time-order built into Einstein’s Special Theory of Relativity. Einstein showed that there is no fact of the matter as to the time order of events in certain cases (namely, when the events are spacelike separated); in these cases, we can only make sense of time order *relative to a state of motion* (“*frame of reference*”).<sup>8</sup> Similarly, the discovery that space is non-Euclidean (built into Einstein’s General Theory of Relativity) shows that we need to relativize the notion of parallelism: two geodesic segments can’t be parallel simpliciter, but only parallel relative to a path of transport from a point of one to a point of the other. This is in some ways a more dramatic example than simultaneity: with simultaneity, it is only events that are very far apart for which relativizing to different states of motion leads to dramatically different simultaneity claims; but for parallelism, even nearby line segments that are highly non-parallel relative to the obvious paths of transport will be parallel relative to others (e.g. ones that go far away and come back, or that loop many times nearby). In these cases, there is little doubt that Einstein’s views went against the views of the person on the street as well as against those of opposing theorists.

Indeed, one might be inclined to say that Einstein changed the meaning of ‘simultaneous’ and ‘parallel’, to frame-dependent and path-dependent notions—though there is some pressure in

---

<sup>8</sup> Here and in what follows, I mean by ‘time order’ *linear* time order: I’m not considering the non-frame-relative partial order given by the light cone structure.

the other direction too.<sup>9</sup> But even if he did, that fact seems of little interest to issues about space and time: the “old meanings” *needed* replacement since there was nothing in reality corresponding to them, and so he replaced them with ones that would serve much the same purposes without the incorrect commitments. Similarly, if we suppose that ordinary practice is committed to a non-relative notion of metaphysical justification, then that “old meaning” needs replacement if it turns out that it is incoherent or that nothing falls under it; we need to replace the old meaning with one that serves the same purposes but without that commitment.

Of course, I haven’t argued that the commitment to metaphysical justification is in fact incorrect, and it isn’t my purpose to do so here. My present interest is just in putting forward a view of justification *compatible with* it being incorrect. The case for it being incorrect is better made once this is done, for if it can be done then the cost of giving up metaphysical justification will be shown to be far less than one might initially have thought—far less than epistemological realists would have you believe.

**3. Evaluation of norms.** The evaluativist view is, as I’ve said, a kind of relativism: in the case of primary interest here, epistemological relativism. But the term ‘relativism’ has had the misfortune of being defined by its opponents.

For one thing, they have typically defined it to mean ‘contextual relativism’: I believe that this would destroy the whole point of the doctrine, and I will have more to say about the far more interesting doctrine of assessor relativism as we proceed.

---

<sup>9</sup> If we do say there has been a change of meaning, say in ‘simultaneous’, what are we to say that the term prior to Einstein “meant”? If it “meant” a relation obeying all the properties that Newton assumed, then since there is no such relation, virtually everything speakers said using the word was false. If it “meant” some two-place relation that is relativistically kosher, then some very surprising claims come out true: for instance, if it “meant” something that coincides with space-like separation, then in Newton’s mouth “There is an event simultaneous both with the birth of Thales and with the birth of Galileo” would have been true. Given this, it may be best to interpret Newton as having used the word ‘simultaneous’ in a way that actually was frame-relative, though he didn’t know it. (Similarly for earlier uses of ‘parallel’.)

Even if we resist this view about Newton, we may want to maintain it for those today who are ignorant of Special Relativity and wouldn’t understand explicitly relativized simultaneity claims. If, as is usually assumed, the meaning of ‘arthritis’ for the lay person is determined by what the experts in his community mean, then it’s natural to suppose that in the mouths of lay people ‘simultaneous’ and ‘parallel’ have a hidden relativity that they don’t know about.



But in addition to this, the opponents have often defined relativism as committed to the idea that all norms are equally good.<sup>10</sup> Certainly no relativism of the sort I want to defend has any such commitment. Important norms, whether moral or epistemological, differ in straightforwardly factual ways that matter to us. A moral norm may be such that a society's mostly following it would lead to vast inequality of resources, or to authoritarianism, or to all manner of other things that we may dislike. Another moral norm may be better in these regards, though perhaps worse in others. If not worse in others, it is better than the first; if sufficiently worse in others, it may be worse overall, or tied. Or more likely, it may be incomparable: 'better than' is presumably only a partial order.<sup>11</sup> Why on earth would anyone claim that any two norms are equally good?

Objection: "These comparative judgements of which norms are better than which are themselves normative, so you haven't really allowed for one entire normative system being better than another." Well, of course judgements of which norms are better than which are normative—what else could a judgement of betterness possibly be? How is it supposed to follow that I haven't allowed for one entire normative system being better than another? Judgements of betterness, even when based in part on non-normative facts, can only be made from norms (taking norms to include systematic preferences). And few norms—none worth taking seriously—will evaluate all norms equally on given factual assumptions.

Perhaps the person who says that relativism declares all norms equally good is defining

---

<sup>10</sup> This is something one often hears in conversation, and there are at least strong implicatures of it in print. For instance, Boghossian (2001 pp. 22, 23) takes relativist and other non-factualist views to imply that no norm is "more correct than any other". Read literally I can't object: since talk of correctness (construed "thickly": see Section 7) doesn't apply at all, then it applies equally. But the formulation certainly seems to *suggest* that the view has it that the norms are equally good; and indeed, Boghossian goes on to say that "there is nothing that epistemically privileges one set of epistemic principles over another" (p.23). [Boghossian informs me that he regards this last quote as careless and not reflective of his views. In his more recent work he continues to say that according to non-factualism, all norms are "equally correct" (e.g. 2006a, pp. 62, 73) and "equally valid" (p. 2), but is careful not to say that they are equally good. He does however give some arguments which in my opinion draw their plausibility from a natural tendency to associate "equally correct" with "equally good".]

<sup>11</sup> For a familiar reason, we can't regard incomparabilities as ties: A can be incomparable to B and also incomparable to C, even though C is clearly better than B; if we were to take incomparability to be a kind of equal goodness, we'd induce intransitivity into the *better than* relation.

“equally good” in a special way, to mean “have a *norm-independent* goodness to equal degree”? But in that perverse sense of ‘equally good’, the relativist will declare everything equally good, since he holds that there is no such thing as norm-independent goodness. I suppose one could say that this makes the anti-relativist complaint all the more powerful: on the relativist doctrine, *everything* is “equally good”. But it is clear that deploying this definition of ‘equally good’ against the relativist is just refusing to take relativism with even an iota of seriousness.

There might be *some* excuse for taking relativism to be committed to the claim that among norms of the same scope, there is none that is *uniquely best*. The rationale would be that it would be hard to make a distinction between a uniquely best norm (of a given scope) and an objectively correct norm. I don’t actually think that this is correct. (An advocate of Special Relativity might argue that a frame of reference in which the center of mass of the universe is at rest is *in some sense* best, while at the same time insisting that it is not objectively correct in the way that Newtonian or Lorentzian mechanics demands.) But in any case, I see no motivation—or rather, no motivation independent of normative realism—for thinking that there is a best norm: it may well be that for each norm there is a better one, and it seems quite likely that there are ties and incomparabilities “all the way up” (i.e. that it isn’t just norms that aren’t sufficiently good that can be equally good or incomparable). It’s true that a normative realist probably has to assume that (among norms of a given scope) one is uniquely best (viz., the “correct” one), but that seems like one of the commitments of normative realism that should be questioned. (Actually, from an evaluativist point of view the question of whether there is a uniquely best norm isn’t straightforwardly factual, since it will be true or false only relative to a norm of goodness. Still, it seems like a pointless constraint on a norm of goodness that it declare one norm of obligation of a given scope objectively better than all others.)

Let’s turn now to epistemic norms. Here the same basic points as above hold, but there is an additional point that requires comment: a kind of quasi-circularity that plays a very important role in meta-epistemology.

First note that in the case of both moral norms and epistemic norms, we can imagine the evaluation of the norms as divided into two stages. In Stage 1 of the evaluation of a given system of morality, I ascertain certain of its straightforwardly factual properties: e.g. (n-i) the

employment of the norm would lead to greater inequality, or (n-ii) it would make promoting the happiness of war-mongering world leaders as important as promoting the happiness of their victims. In Stage 2 I make moral evaluations of these non-moral properties. (This doesn't really need to be considered a later stage; the point is really that it involves conditional evaluations, evaluations conditional on an assumption of the straightforwardly factual properties.) Similarly in epistemology: In Stage 1 I ascertain that a given normative system (say an inductive method) (e-i) would lead to rapid learning from experience in circumstances X but not in circumstances Y (where "learning" is construed non-normatively, as just modification in one's dispositions to predict), or (e-ii) doesn't allow for the relevance of certain kinds of information. In Stage 2 I make an evaluation of the system on the basis of these non-normative properties (perhaps in comparison to the non-normative properties of other normative systems): e.g. I declare that because of the features enumerated, the system is not a good one, and that employing it is unjustified or irrational or whatever. So there is a structural similarity to the moral case (and claims that "all systems are equally good" seems as absurd in the epistemic case as in the moral).

There is another point of commonality between the normative and epistemic cases, which leads however to a salient difference. These involve how the non-normative facts are gathered at Stage 1. The commonality is that in both cases, ascertaining the non-normative facts ((n-i) and (n-ii) or (e-i) and (e-ii)) would require the use of ordinary empirical and logical reasoning, but not normative reasoning. The resultant difference is that in the epistemological case this may seem somehow circular: the very methods that are being evaluated must be employed at Stage 1 of the evaluation procedure. This can seem more disturbing than the use of moral or epistemic norms at Stage 2 of the evaluation, which occurs in both the moral and epistemic case.

The sense of circularity here will need some discussion. There is certainly no problem about how "the facts together with the norms" determine the evaluations. If there is a problem, it is a problem about how *we* determine the evaluations. But that isn't much of a problem either: by and large, we determine the evaluations by following the epistemic methods or norms that we normally follow. The real concerns here are about the significance that these evaluations have.

Two such concerns are *the problem of immodesty* and *the problem of modesty* (both named in reference to Lewis 1971 and 1974). The *problem of immodesty* is that it may seem that

any method will positively evaluate itself, in which case positive self-evaluation doesn't really cut any ice. The *problem of modesty* is that if this isn't so, then some methods will negatively evaluate themselves, and presumably recommend other methods over them; but then we have a situation where a method tells us not to follow itself, which seems somehow incoherent. (Or at any rate, such a method seems not to be consistently followable, since consistently following it would require not following it.) These initial remarks about the problems will need refining (see Section 11), but I hope they are enough to indicate that there are quasi-circularity worries about epistemological evaluation that don't arise in the moral case, and they are connected with some of the central questions of epistemology.

But the present issue isn't the importance of the quasi-circularity, but whether it in combination with relativism has relevance to the claim that all epistemological norms are equally good. And it is hard to see that it does have relevance to that: as above, the question of whether one norm is better than another is obviously a normative claim, and so according to the relativist can be assessed only in a norm-relative way; and no system of norms that anyone would take seriously will imply that all epistemic norms are equally good. The fact that this judgement is norm-dependent in a deeper way in the epistemological case than in the moral is interesting, but doesn't affect this basic point.

**4. Norms.** Boghossian 2006b calls attention to an important ambiguity in the term 'norm'. On the *propositional* use, a norm is a fairly general normative proposition. If norms are so conceived, it is easy to explain norm-relative properties: e.g., X ought to do A relative to norm N if and only if N and the non-normative facts together entail that X ought to do A. But on this conception of norms, there is a *prima facie* problem with explaining what it is to *accept* a norm: it can't be to *believe* the norm, for then only normative realists could accept norms. An example of a *non-propositional* use would be to take norms as conditional imperatives: if one is in circumstances C, do P! On this use, there seems to be no problem in understanding what it is to accept a norm, but there is a *prima facie* problem in explaining norm-relative properties.

So on either use there is a *prima facie* problem to be overcome, and vacillation in the use of the term could lead the relativist to bury difficulties under the rug. But it doesn't really matter very much which of the two uses the relativist picks: the same problems arise in either case, in

different guises.

Suppose norms are conceived as fairly general normative propositions. Then what should the relativist say about what it is to accept such a norm? I think that will depend on the kind of normative proposition in question: whether it is for instance a normative proposition about what is *good* or one about what one *ought to do*. The idea should be that acceptance of a propositional norm *about what's good* is to be explained in terms of one's preferences, or of the preferences that one approves of; and that acceptance of a propositional norm *about what one ought to do* is to be explained in terms of one's policies, or of the policies that one approves of. (If there are propositional norms expressed in terms of normative notions not explicable in terms of 'good' and 'ought', we'll need other mental states than preferences and policies; I won't attempt to catalog them.)

I myself will adopt the other, non-propositional, tack:<sup>12</sup> I will take the norms to simply be the preferences and policies (and any other items that go into the final catalog). The preferences will include "higher order" preferences: preferences about preferences, and preferences about policies. This is so that both approval of preferences and approval of policies can get into the discussion. I'll be concentrating mostly on 'ought' judgements, so until the time when it is important to consider higher order judgements about which norms are good, it is policies rather than preferences that I will focus on.

So one kind of norm is a policy: in the moral case, a policy both for acting and for gathering information on the basis of which to act, and in the epistemic case a policy both for believing (or believing to a certain degree) and for acting so as to improve one's epistemic situation (e.g. by trying to gather more evidence, or to think up more possible explanations, or to determine whether an answer to a question of interest follows from things one already accepts). Policies are sometimes stated in normative language ("You shouldn't believe a conjunction without believing the conjuncts"), but here the normative claims are generated by the policy: in the example, the policy is something like an imperative ("Don't believe a conjunction without

---

<sup>12</sup> In doing this I am influenced by the opening chapter of Gibbard 2003, which takes norms of rationality to be *plans*. Among the advantages of the non-propositional choice is flexibility: it makes it easier for, e.g., both the policies one follows and the policies one approves of to get into the discussion.

believing the conjuncts”), and the “shouldn’t” formulation just means that if you act in the way suggested you are violating the policy.<sup>13</sup>

Of course, many policies for acting (or gathering information on which to act) aren’t moral policies. For instance, policies for making money aren’t moral policies except in exceptional circumstances. Only moral policies should count as moral norms. But what makes a policy a moral policy? I’m not in a position to give any answer that isn’t embarrassingly superficial, but for the most part I know one when I see one (as do you)—or when I (and you) don’t, it’s usually because the boundaries of the moral are quite vague. In any case, I don’t see that there’s any special problem for the evaluativist here: presumably any answer that the realist gives as to how moral *oughts* differ from other *oughts* can be adapted to an answer to how moral *policies* differ from other *policies*.

Similarly in the epistemological case: only epistemological policies count as epistemological norms. For instance, a policy of popping pills to make one believe that all is well with the world doesn’t count as epistemological, and this may have something to do with its not being done in pursuit of truth; but I wouldn’t attempt any precise account of the epistemological. Indeed, I suspect that there is no sharp distinction between the epistemological and the merely pragmatic: the pragmatic need for computationally tractable theories shapes our choice of theories in many ways, but where the decision is counted *merely* pragmatic and where epistemological strikes me as rather arbitrary.

One obvious kind of question to raise about policies and other norms is what kind of role

---

<sup>13</sup> As Boghossian 2006b emphasizes, there are norms of explicit permission as well as of obligation or prohibition: a logically omniscient being could read the permissions from the obligations or prohibitions ( $p$  is permitted if it is not prohibited, i.e. if *not- $p$*  is not required), but those of us who aren’t logically omniscient can’t compute what does and doesn’t follow from our explicit obligations, so we need a system with both explicit obligations and explicit permissions. (This introduces dangers of conflict: the system could both explicitly permit  $p$  and implicitly prohibit  $p$ , thus being practically inconsistent. But such is life.) One way to understand norms of explicit permission is as explicit rejection of policies (so that a norm isn’t really a policy, but the acceptance or rejection of a policy). Alternatively, the notion of policy could be broadened: say, as a partial function that assigns to certain pairs of input conditions and possible actions either ‘+’ (for explicit permission) or ‘-’ (for explicit prohibition); when there is no mark, permissions and prohibitions (and obligations, i.e. prohibitions of the opposite) can still be implicit in the other explicit permissions and prohibitions, but aren’t explicit. For simplicity however, I will ignore explicit norms of permission in what follows.

they play in the lives of people. I think there is no single role that they play. For instance, a person can *act or believe in accordance with* a norm (or largely in accordance with it), or her acting and believing may in some sense *guided by* the norm. Or she can *make her evaluations in accordance with* a norm (or largely in accordance with it), or in a way that is guided by the norm. Or she can *be committed to* acting or believing in accord with a norm (or have a high degree of commitment to doing so), or to evaluating in accordance with it. The norms one bears these relations to needn't coincide. (And the relations can themselves subdivide: e.g., one can employ different norms in evaluating oneself than in evaluating others.) While it is natural to talk of “an agent's norms”, and to ask such questions as “how an agent's norms evolve”, such talk can be misleading because it does not differentiate between the different ways in which an agent can be related to a norm. Indeed, I think a key part of the story of how the norms that the agent is related to in various ways evolve is going to involve the interactions among these different relations to norms.

Even confining attention to a single way in which an agent is related to a norm, there will be many different norms or policies to which an agent is related. Some are very low level. My actions might for instance be guided by the policy of generally believing what I read in the NY Times, unless it appears under the byline of one of a small number of reporters who I know to be shills of the Bush administration. There are many similar such “low-level” policies that guide my behavior. These are policies that I can easily revise as I gain more information. How do I make the revision? By following a broadly inductive policy. This is another epistemic policy; intuitively, a “higher level” one. One possible view is that there is a “highest level” norm or policy that guides my behavior at any time. (This view is advocated, for instance, in Pollock and Cruz 1999; indeed, they advocate an extreme form of it, according to which this “highest level” norm can't be rationally revised.) I do not presuppose this—in fact, it strikes me as quite a dubious supposition, and I think that evaluativism helps to undermine some of the pressure to make the supposition. (There will be a brief discussion of this in Section 12.) I will however put most of my focus on relatively “high-level” norms; indeed, you can take my discussion to be confined to policies that could serve as a “highest level norm” for guiding a person's behavior, if you believe there to be such a thing.

Putting aside these questions about how people are related to norms, let's turn to some questions that arise for norms in abstraction from their relation to agents: questions about, broadly speaking, the logic of normativity.

At least for norms that are policies, the idea of an action, belief, etc. being reasonable-relative-to-a-policy admits an obvious explication in non-normative terms: it means that acting, believing etc. in the manner in question, given the circumstances in which the agent finds herself, is compatible with the policy.<sup>14</sup> An alternative explication would be possible: that acting, believing etc. in the manner in question, given the circumstances in which the agent finds herself, is *dictated* by the policy; that is, it would be *incompatible* with the policy *not* to believe it in those circumstances. This latter explication strikes me as unnatural, though. It might be more natural for 'justified' than for 'reasonable'—this is another way (beyond the one mentioned in note 2) in which the two terms might be thought to diverge in ordinary usage. But rather than using divergent terms, I'll stick to 'reasonable', and talk of *reasonably failing to act or believe* (relative to a policy) as well as *reasonably acting or believing* (relative to a policy).

So we have a norm-relative notion of reasonability, explicated (when the norms are just policies) in non-normative terms. This is an *all-or-nothing* notion of reasonableness, which may not be ultimately what we want: a graded notion, in which there is a partial ordering of degrees of reasonableness, may be more useful, and this might require complicating our picture of what a norm is. But I'll make do with the simple notion here.

**5. Norm-relative truth.** A norm-relative notion of reasonableness induces a norm-relative notion of truth for sentences about reasonableness. To see this, first consider ordinary context-relative claims where the relativity is not explicit, e.g. 'It is raining' (with its hidden contextual relativity to a spatio-temporal location). It is possible to apply the term 'true' not just to individual tokens of 'It is raining', but to the sentence type. If we do so, then the nature of truth guarantees that

- (1) 'It is raining' is true if and only if it is raining;

---

<sup>14</sup> If you think that compatibility is itself normative, you can replace it by compatibility in a given system; this can be explained in syntactic or model-theoretic terms, which certainly aren't normative.



what this amounts to, given the hidden relativity, is (the universal generalization of)

‘It is raining’ is true at spatio-temporal location  $l$  if and only if it is raining at spatio-temporal location  $l$ .

In other words, the truth attribution

(1L) ‘It is raining’ is true

in (1) itself has a hidden contextual relativity. This is to be expected: the nature of truth guarantees the equivalence between (1L) and

(1R) It is raining;

the last has a hidden relativity, so the former must as well.

This application of ‘true’ to sentence-types is even more natural in connection with controversial cases of relativity. Consider

(2R) The birth happened just before the star exploded.

To those wholly ignorant of 20<sup>th</sup> century physics, this seems to have no relativity (beyond any in ‘the birth’ and ‘the star’, which I will ignore); so that when they say

(2) ‘The birth happened just before the star exploded’ is true if and only if the birth happened just before the star exploded,

they do not require any relativity in the truth predicate. Those who know of Special Relativity will still make utterances of form (2R), but will “intend them as” having a hidden relativity to a state of motion in the predicate ‘before’.<sup>15</sup> Similarly, they will still make utterances of form (2), or its left hand side

(2L) ‘The birth happened just before the star exploded’ is true,

but will intend these as relative too, this time with a hidden relativity to a state of motion in the predicate ‘true’ (as well as in the ‘before’ on the right hand side of (2)). So the claim (2) should be construed as

(For all states of motion  $f$ .) ‘The birth happened just before the star exploded’ is

---

<sup>15</sup> I put ‘intend them as’ in quotes because there is some issue about how to understand it. Certainly in a typical utterances one does not explicitly think about the relativity in ‘simultaneous’, ‘prior to’, ‘energy’, ‘momentum’, ‘parallel to’ and all the other terms one knows involve some sort of relativity. Perhaps the best thing to say isn’t that one “intends” the relativity, but that one has a standing view that when one makes such utterances or thinks such thoughts, a relativized interpretation is appropriate. But I’ll continue to speak in the simpler way.

true relative to  $f$  if and only if, relative to  $f$ , the birth happened just before the star exploded.

At least, this is so when the quotation mark name in (2) or (2L) is taken to refer to the semantic type of an utterance of it by someone who accepts Special Relativity. But how broadly should one construe that semantic type? Should one construe it as including utterances of the orthographic type (2R) made by speakers who don't recognize the relativity in temporal notions? Indeed, there are three cases of this to consider: we must consider utterances of (2R) made by

(i) someone who lived before Relativity Theory;

(ii) someone today who is ignorant of Relativity Theory;

and (iii) someone today who knows of it but rejects it.

Which if any of these count as having the same semantic type as the utterances of (2R) made by those who accept the relativity? This raises controversial issues in the philosophy of language. For instance, many will want to say that at least in case (i), the speaker should be understood to have meant "absolutely before", so that all their utterances involving 'before' were false; in this case, there is no relativity in applications of 'true' to *their* utterances of time-order. (This line can be taken in cases (ii) and (iii) as well, though there it seems to be in some tension with widely-held views about the semantic features of 'arthritis' in the language of those with bad misconceptions about its nature.) My own view, for what it's worth, is that for theories that are overall false, there is often no determinate fact of the matter as to the semantics of individual sentences; and this is so in particular in the case of theories that fail to recognize important relativizations. Even when it's clear that the theorist does not recognize the relativization in a predicate, as in (i)-(iii) above, it may be appropriate in some contexts to treat a given sentence involving the predicate as making a relativized claim and in others as making an unrelativized claim. And it isn't that one way of treating it is right and the other wrong: there is no determinate fact of the matter. That's my own view, but there is no need to decide on these issues here. The only thing I want to take a firm stand on is that whatever hidden relativity you posit in any utterance of (2R), it should get carried over into the predicate 'true' in the corresponding utterances of (2L), i.e. the utterances of (2L) in which the quotation-mark name refers to the semantic type of the utterance of (2R).

The relativity I postulate in evaluative terms is a kind of controversial relativity: it is

controverted at least by many theorists of evaluative discourse. (Whether it is opposed to the views of the person on the street is a more delicate question: such a person would certainly not be aware of any theoretical articulation of it, but I suspect that thoughtful people have far more inklings that there there is a relativity here than pre-Einsteineans had about a relativity in time-order.) There is, as I've said, a crucial difference between the evaluative case and the time-order case: in the time-order case we have ordinary contextual relativity, whereas in the evaluative case we have a relativity of a very different sort, that will be elaborated in Section 10. (A consequence of the difference is that in the norm case, typical ways of making the relativity explicit destroy the evaluative nature of the utterance.) But despite this hugely important difference, what I've said for contextual relativity carries over: it's perfectly appropriate for an evaluative relativist to say

(3) 'That belief is reasonable' is true if and only if that belief is reasonable, but this must be understood as involving a hidden relativity to a norm, both in the unquoted 'reasonable' on the right and in the 'true' on the left.

One possible difference with MacFarlane is that he *seems* to advocate a kind of relativism in which the relativity of the truth predicate doesn't derive from a relativity in the ground level propositions to which it applies. If that really is his view, then I don't understand it; but I think that what I'm saying is not far from the spirit of his views.

Again, my claim that there's a hidden relativity in 'reasonable' applies *at least* to attributions of truth to sentences involving 'reasonable' *as used by relativists*. How one wants to apply 'true' to utterances involving 'reasonable' in the mouths of (i) committed anti-relativists, or (ii) the person on the street, involves issues like those in the simultaneity case: one could treat these utterances either as relative, or as mostly false because nothing falls under their predicate, or perhaps in some other way. My own view is that there is no determinate fact of the matter on how to treat them, but this is not central to my overall position on normativity and I will not argue the matter here.

**6. Pure and impure degrees of belief.** A policy, as I've explained it, prohibits certain actions, beliefs or whatever, in certain circumstances: if you like, it consists of a set of conditional prohibitions. They are prohibitions in the sense that acting or believing in violation of them is

inconsistent with carrying out the policy. Let's restrict attention for the moment to policies that are completely precise. Then a policy divides up an agent's possible activities of acting or believing in certain ways, and of not acting or not believing in those ways, into two classes: those prohibited and those not prohibited. (Indirectly, we get another division, those demanded and those not: for the set of "possible activities" as I've defined it is closed under negation, and  $A$  is demanded when its negation is prohibited.) These divisions are dependent on the state of the world (and an agent's location in it): a policy may prohibit some activity  $A$  relative to one possible world  $w$  in which the agent may find himself but not relative to another.<sup>16</sup> Policies are special cases of what Gibbard calls *complete norms*. The term 'complete' has connotations that might be misleading. In particular, (i) "complete" norms might be taken to be norms for which there is no gap between what's permitted and what's demanded, but Gibbard doesn't assume completeness in that sense; (ii) "complete" norms might be thought to be ones with a kind of maximal scope, but even very narrowly focused norms can be "complete" in the sense Gibbard specifies. So I will use the term 'precise norm'.

A precise (or "complete") norm  $n$  is (by Gibbard's definition) something that in conjunction with any possible world  $w$  determines a truth-value for every evaluative statement  $A$ . Policies as defined above meet this condition: e.g. if the norm prohibits believing  $p$  in circumstances  $C$  but in no other circumstances, it determines the value *true* for 'You should refrain from believing  $p$ ' in those worlds in which  $C$  is true and the value *false* in the others.<sup>17</sup> Restating this slightly, the precise norm  $n$  in effect assigns to each evaluative statement  $A$  a function  $/A/n$  from possible worlds to truth-values. I'll call such a function a Stalnaker-proposition (though this might not be strictly appropriate, given note 16). Indeed, we can regard the precise norm as assigning a Stalnaker-proposition to every statement  $A$ , evaluative or not: if  $A$  is not evaluative, then every precise norm assigns it the same Stalnaker-proposition, viz. the

---

<sup>16</sup> I'll take a "possible world" to include the spatio-temporal location of the agent in the world: that is, it is to be what Lewis 1986 calls a *centered* possible world.

<sup>17</sup> Of course, the worlds at which  $n$  declares 'You should refrain from believing  $p$ ' false (not true) will not typically be the same as the ones in which it declares 'You should believe  $p$ ' true: it will declare both false if it neither prohibits nor demands belief in  $p$ , and it will declare both true if it both prohibits and demands belief in  $p$ . (I haven't required that norms be consistent, much less that they be "consistent relative to the facts of each possible world".)

function  $|A|$  that gives its ordinary truth value in each world.<sup>18</sup>

It will be useful below to weaken the requirement that policies and other norms be precise. A simple-minded way to do so, adequate to present purposes, is to view an imprecise norm or policy as a set  $N$  of two or more precise norms or policies: if something is permitted by some member of  $N$  but not by another, then  $N$  leaves it “fuzzy” whether that thing is permitted. (Similarly for demanded.)

Now suppose we represent a person  $S$ 's degrees of belief in non-evaluative propositions by a probability measure  $\mu$  over (a sufficiently rich  $\sigma$ -algebra on) the set of possible worlds. This determines a probability function  $P$  on non-evaluative claims:  $P(A) = \mu(|A|)$ , where  $A$  is non-evaluative. More generally, it determines a function  $P^*$  that assigns to each precise norm  $n$  a probability function  $P^*_n$  on all claims, evaluative and non-evaluative:  $P^*_n(A)$  is just  $\mu(|A/n|)$ . (Note that—contrary to what Gibbard's formulation of expressivism might suggest—the only probability measure employed here is over the set of worlds, not over the set of norm-world pairs.)  $P$  assigns to each non-evaluative claim  $S$ 's *pure* degree of belief, unmixed by normative evaluation. If  $S$  is committed to the precise norm  $n$ ,  $P^*_n$  extends  $P$  by assigning to each evaluative claim a real number (from 0 to 1) that gives  $S$ 's *impure* degree of belief in that claim—impure because it contains the evaluative element given by  $n$  in addition to the pure belief component given by  $P$ . (On non-evaluative claims,  $P^*_n$  agrees with  $P$ .)<sup>19</sup> So for instance if

---

<sup>18</sup> I don't mean to take a stand on how easy it is to find statements in English with no evaluative elements: it's compatible with what I say here that most or even all apparently descriptive predicates are to some extent “thick” with evaluation. I do hold, with Blackburn and many others, that it is important to disentangle the descriptive and evaluative elements in “thick” predicates: indeed, the excellent discussion of ‘cute’ in Section 4.4 of Blackburn 1998 shows that failing to do so undermines the possibility of normative critique of entrenched views.

<sup>19</sup> I'm being a little loose here, because a person can be committed to many policies at the same time; if we want to define a unique notion of degree of belief for evaluative claims on this model, we need to single out one of them. This will presumably involve (i) figuring out which one or ones are fundamental enough to take as contributing to the desired notion of degree of belief, and (ii) if there is more than one that is fundamental enough, moving to the “union” of them, i.e. the policy that contains the prohibitions in each of them. (Of course, the union may be inconsistent even when the individual members are consistent, but I've observed in note 17 that the model under discussion does not require that policies be consistent.) It is also possible to let the policies that count as “fundamental enough” be a matter of context, in which case the notion of impure degree of belief will be as well.

One could define variant notions of impure degree of belief by using, say, the norm an agent uses in evaluating herself rather than the norm she is committed to; but I think that the

the norm  $n$  that we're imagining  $S$  to be committed to demands that an agent believe a certain perceptual claim  $p$  if and only if he satisfies some naturalistic condition  $C$ , then  $S$ 's impure degree of belief that Jones *ought to believe that*  $p$  will coincide with  $S$ 's degree of pure belief that Jones is in conditions  $C$ . If the person employs only an imprecise norm  $N$  (obviously a more realistic assumption), then any probability function  $P$  over non-evaluative claims generates a *non-empty set of* probability functions over evaluative claims,  $\{P^*_n \mid n \in N\}$ : she will not have point-valued impure degrees of belief, even if her pure degrees of belief are point-valued. (Of course, if we want to be at all realistic, we won't assume that even the pure degrees of belief are point-valued.)

There is a verbal question, on this view, of whether impure belief should count as "belief". That is, suppose an agent is committed to norm  $n$ , and has degrees of belief in non-evaluative claims given by the function  $P$ . Should we think of the extended function  $P^*_n$  as literally giving his degrees of *belief* in arbitrary claims, properly so called? Or should we think of it as giving degrees of *something formally like beliefs but not the real thing*. My view is that there's no "properly so-called" about it: we can reasonably talk either way. Perhaps the best course is to avoid taking a stand, by the means employed above: use 'degree of pure belief' for the degrees of belief in non-evaluative claims, and 'degree of impure belief' for the norm-dependent extension.<sup>20</sup> The question of whether this marks a distinction between a special kind of degree of belief and degree of belief generally, or between degree of belief and something formally like it, is a matter for verbal legislation. We'll see though that there is a naturalness to talk of impure belief: it connects up more directly with the phenomenon of normative disagreement.

**7. "Not straightforwardly factual".** I've described evaluativism as a view according to which statements of justification "aren't straightforwardly factual". By this, what I mean is that they

---

resultant notions are less natural and less likely to be confused with a kind of pure belief.

<sup>20</sup> Actually it's probably better to mark the distinction between pure and impure slightly differently: instead of regarding the degree of belief in "If he's justified in believing that  $q$  then he's justified in believing that  $q$ " (or the conjunction of this and some straightforwardly factual claim) as impure because it is a degree of belief in a claim that contains an evaluative term, we can regard it as pure because the claim in question is *norm-insensitive*—the norm drops out of the determination of the degree of belief. But I won't bother to distinguish between the evaluative and the norm-sensitive in what follows.

have a hidden relativity, somewhat analogous to the hidden relativity in ‘simultaneous’. As I’ve said, there’s a very important disanalogy: in the latter case, the relativity is naturally viewed as contextual, whereas in the normative case it is not. (One could come up with fanciful stories on which the relativity in ‘simultaneous’ behaves more like that in normative discourse—e.g. where there is pressure for all speakers to adopt a common frame of reference, and where speakers use not-explicitly-relativized simultaneity claims that are correct only in the frame of reference they prefer in order to pressure others to adopt that or a similar frame—but it is not worth pursuing this.)<sup>21</sup> Even in the Special Relativity case, there is some naturalness in the claim that not-explicitly-relativized statements (e.g. of form ‘*b* is simultaneous with *c*’) aren’t straightforwardly factual, but only factual relative to a frame of reference. I take this to be more natural still in the case where the relativity is assessor-relativity. To call a claim “straightforwardly factual” is not intended as a positive characterization of it, but simply as a denial that it has any of the characteristics that would make it appropriate to call it less than straightforwardly factual. In the present context the only such characteristic that is relevant is assessor relativity, so you can if you like replace ‘straightforwardly factual’ by ‘factual in a non-assessor-relative way’.

But calling statements about justification (or other assessor-relative statements) “not straightforwardly factual” could mislead. For one thing, it could be taken to imply that justification claims aren’t “truth-apt”: that is, that the term ‘true’ is never to be applied to them. This is certainly not my view: in my view, the word ‘true’ has an important logical role that is as important for normative claims as for others. If someone expresses an elaborate normative theory, each part of which seems acceptable but which has a normative conclusion I strongly dislike, I may express my own normative attitude toward it by saying “Not all of his theory can be true, though I haven’t yet figured out which claim in it isn’t true”. I can do this because I take “‘*p*’ is true” to be equivalent to “*p*” in a fairly strong sense: in particular, *any assessor-relativity in “p” is inherited into “p’ is true”*.

---

<sup>21</sup> Indeed, some terms thought of simply as having contextual sensitivity do behave in some ways like the way I’m claiming normative terms behave. This is especially clear in vague terms like ‘rich’: while they admit of standard contextual relativity, there also seems to be a normative element as to where one places the border. (“You think having \$5,000,000 makes you rich? What about having \$100,000,000—now *that’s* rich.”) See Egan forthcoming.

That's my preferred use of 'true' (and I think it should be everyone's), but what about 'fact'? There is something to be said for the view that to say "it's a fact that  $p$ " is also just a pleonastic way of saying " $p$ " (though in ordinary English 'fact' does not seem to have the generalizing role that motivates the view in the case of 'true'). If that's how one prefers to use 'fact', then if one makes a normative judgement (e.g. that skepticism is unjustified), one will equally judge that *it is a fact that* skepticism is unjustified. Still, this will not deny that there is assessor-relativity in "Skepticism is unjustified", it is simply that that assessor-relativity is inherited into "it is a fact that skepticism is unjustified".

There is a question of how far it is reasonable to push this line. An extreme application would be to take the same line about 'it is a straightforward fact that  $p$ '. Were I to adopt that line, I would judge that it is a *straightforward* fact that skepticism is unjustified, even though the claim is assessor relative. But it doesn't seem useful to adopt that line: it would make the term 'straightforward' redundant. It would be roughly analogous to a supervenient saying of a borderline case  $S$  of baldness that  $S$  must be either determinately bald or not determinately bald, on the ground that in each precise valuation  $S$  is either determinately bald or not determinately bald. 'It is a straightforward fact that', like 'determinately' in the supervenient picture, is supposed to be a strengthening of 'true' that give one the means of commenting on the status of the claims in its scope.<sup>22</sup>

Of course, this assumes that the strengthenings make sense. In both the cases of 'determinately' and 'straightforwardly', we need an account of just what the strengthening comes to. In the case of 'straightforwardly' I have suggested that a crucial component is absence of implicit relativity, and in particular absence of assessor-relativity; more particularly still, the absence of sensitivity to policies, preferences, etc. (i.e. to norms as I'm understanding them), which is a kind of assessor-relativity but possibly not the only kind. The difference between statements that are norm-sensitive and those that aren't ultimately comes to the difference between the conceptual and social roles of norms on the one hand and pure beliefs on the other.

---

<sup>22</sup> There is a similar issue about 'property'. Like 'true', this can be given a pleonastic use, on which any meaningful predicate without contextual relativity corresponds to a property. Given this, I wouldn't want to say that normative predicates don't express properties; I say instead that they don't express straightforward properties.



Perhaps there is room for skepticism about whether such an account can be provided in the end. I've allowed the relativist to assert normative claims, to assert that they're true, even to assert that they state facts. Indeed, I haven't prohibited her from saying that they express "straightforward facts", though I've said that this would be highly misleading and that it's hard to see why she would want to say that. The worry is that if I allow her to say such things, haven't I given the factualist everything he should want? (This is the "creeping minimalism" worry of Dreier 2004.)

There are two responses to this worry. One relies on the fact, stressed before, that an evaluativist needn't accept, and probably shouldn't accept, that any one norm is best. The factualist needs to hold that one norm is objectively privileged, and the non-factualist needs to deny it. The creeping minimalism worry concerns whether one can make sense of the issue; but if the evaluativist holds to standards on which no norm is even *best*, that position would seem to suffice for non-factualism.

For a second response, available even if for some reason the evaluativist holds some norm to be uniquely best, consider again relativity in physics. Here too, I as an advocate of relativity about temporal order don't rule out the use of ordinary unrelativized language: most of the time, I don't talk in an explicitly relativized way, and I don't think in an explicitly relativized way either. (This is even more obvious in the parallelism example of Section 2). Not only do I make unrelativized (i.e. not explicitly relativized) claims about temporal order, parallelism, etc. in my speaking and writing and thinking, I regard such claims as true. Those truth attributions have the same "implicit relativity" as the ground level claims, but here too I usually don't explicitly think the relativity. What then distinguishes me from the person who doesn't accept Relativity Theory? After all, in both cases our explicit thoughts and utterances are of such forms as

A is (nearly) simultaneous to B

and

'A is (nearly) simultaneous to B' is true.

I think that the answer to this is that while in many contexts I may on a superficial level talk and think just like those ignorant of Relativity, I do not do so when the chips are down:

when it matters, I explicitly relativize. And that's basically so in the normativity case too: in certain contexts of persistent disagreement about whether I ought to believe X, I will back off the norm-sensitive language, and say instead something like: "Relative to such and such standards I ought to believe X; moreover, I advocate those standards, because they have such and such properties, which I strongly favor". (Note that the way I back off is a bit more complicated in the normative case, because the relativity is assessor-relativity.) I think we could in principle conduct all normative debate in this sort of terminology, where we fully disentangle the impersonal straightforwardly factual aspects (what I should do relative to such and such standards, and what straightforwardly factual properties those standards have) from the autobiographical factual properties of what my normative attitudes are (what standards I advocate, what properties of standards I prefer). Or rather, we could do so if we knew enough of the details of our own norms and those of the other participants in the debate, and also had the super-human logical skills to trace out all the logical implications of factual beliefs on our own normative attitudes and on those of the other participants. But given that we don't have such knowledge, it would be highly impractical to eliminate normative language in most circumstances, though there are circumstances in which it is helpful to do so.

**8. Norms vs. worlds: an untenable dualism?** It is clearly part of the evaluativist picture I'm suggesting that "worlds" determine the truth-values only of *non-evaluative* claims; because "normative facts" aren't included in the worlds, then a world *plus a norm* is required in order to generate truth values for evaluative claims. But if we were to assume normative realism, then the worlds themselves would include the normative elements needed to generate truth values for evaluative claims; in that case the probability measure  $\mu$  should be extended to these enriched worlds, and the norms would be unnecessary for this purpose.

This raises the question of whether the normative realist should accept the distinction between norms and worlds that is at the basis of evaluativism as I've explained it. If the worlds contain normative facts, what room is there for norms?

But I think that worry would be a mistake: norms, in the sense that I and Gibbard have employed the term, are just things like policies and preferences, and these are distinct from worlds even if worlds include normative facts. So on a normative realist picture, there will still

be norms in Gibbard's sense, and they will play many of the same psychological roles in explaining behavior that they do in the evaluativist picture. Norms in this sense don't say anything about the world, even if the world includes normative facts. What's distinctive about normative realism isn't that norms (in the Gibbard sense) are just pure beliefs of a special sort (viz., pure beliefs about the normative facts: "pure normative beliefs"); they aren't. Rather, what's distinctive is that according to normative realism there is a special relation between norms (policies) and pure normative beliefs. The special relation arises because the normative realist has a notion of a norm (policy) being *straightforwardly correct*: it is straightforwardly correct if the way it dictates we act or not act agrees with the straightforward normative facts about how we should act or not act, i.e. with the normative element that goes into the worlds. So a normative realist will have both norms in the sense of policies, *and also* pure normative beliefs; this means that an agent who is a normative realist will have both impure normative beliefs and pure ones. But the impure and the pure will be hard to distinguish: they will extensionally coincide, given that the norm being relativized to in the definition of impure beliefs is one that the agent is committed to, which will presumably coincide with what he believes to be objectively correct. Because of this, the normative realist may be inclined to attribute some of the psychological roles that the evaluativist attributes to norms to the normative beliefs instead.

I haven't tried to argue in favor of the evaluativist view, but only to provide an initial sketch of it. (Further clarifications will emerge.) It is clear that there is a challenge that the evaluativist needs to meet: he must set out a believable picture of how normative discourse works that does not require any straightforwardly normative properties or any belief in them. But the normative realist has his own challenges. One of these is to give the belief in normative facts a believable role. Why make our policies conditional on our beliefs about the existence and nature of normative facts? If we morally disapprove of torturing dogs, why rest this disapproval on a pure belief that there is a straightforward normative fact that we oughtn't torture dogs?

Indeed, I'm tempted to say that the moral realist has not only a dubious metaphysics, but also a dubious morality that allows torturing dogs under the condition that there are no straightforward moral facts, or under the condition that those moral facts permit or even require such torture. (While the believer in straightforward moral facts is likely to bend his opinion

about what those facts dictate to his antecedent moral judgements, this needn't be so, especially among those who defer to authorities for their opinions about the alleged moral facts.)<sup>23</sup>

Similarly, the epistemological realist has not only a dubious metaphysics, but also a dubious epistemology that allows believing a conjunction without believing the conjuncts under the condition that there are no straightforward epistemological facts. The realist view seems to entail that if one allows any degree of doubt at all about the existence of normative facts, it's reasonable to weaken one's commitment to conjunction elimination accordingly.

**9. Embedding, and a first stab at normative debate.** I've been discussing issue (I) from Section 1, but the answer to it makes clear that there is no real problem over issue (II): evaluative terms will embed in a perfectly straightforward way. (The point was well made by Gibbard; as noted earlier, the current view seems to be a notational variant of the central core of Gibbard's view.) For instance, in the sentence

If believing that the Earth is round is justified, then believing that people on the other side will fall off is also justified,

there is a single hidden parameter that occurs in both occurrences of 'justified'; someone who adheres to a precise norm  $n$  will evaluate this sentence according to that norm together with his beliefs about the world, and if those beliefs together with his norm make it unjustified to believe that the Earth itself gravitationally attracts then the evaluation of the above claim might well be positive. Of course, in this case it's hard to find plausible norms which would have this result *given the non-evaluative facts*: anyone who asserts the above sentence is almost certain to be making a straightforward error.

Issue (III), concerning normative debate, is far more complicated. I mentioned that there are several different important relations between norms and agents: e.g. there are the norms that

---

<sup>23</sup> The point made here is influenced by Dworkin 1996 and Street forthcoming; both say that the issue of realism is a normative doctrine. (I think they, or at least Dworkin, means that it is *solely* a normative doctrine: that there is no metaphysical issue, only the normative. This I do *not* accept.) I'm also influenced by conversations with Melis Erdur.

I don't believe the point is adequately answered by noting that an erstwhile realist who discovered there to be no straightforward normative facts would almost certainly continue with those moral preferences, and give up her view that she must base moral preferences on beliefs about such facts. For it's still the case that *while she's a realist* she has those conditional preferences; and those conditional preferences strike me as morally objectionable.

an agent *employs in making evaluations*, those that he *employs in acting*, and those that he is *committed to* employing in these ways. The most interesting issues about normative debate involve how debate can lead to changes in the norms an agent is committed to and in the norms he actually employs. This will turn out to be intimately connected with another important issue, which is how an agent resolves inconsistencies within a given norm or between different norms.<sup>24</sup> (The ‘between different norms’ case includes cases of inconsistencies between the dictates of two norms he is committed to, cases of inconsistencies between the dictates of two norms he employs, and cases of inconsistencies between the dictates of a norm he is committed to and of a norm he employs.)

But let us put aside these hard questions for now, by some ridiculous oversimplification. Let’s imagine an argument about the reasonableness of a particular action or belief, among agents each of whom is committed to a single relevant precise norm. (I do not assume that it is the same norm for the different agents.) Imagine also that each agent’s norm will remain unchanged during the debate, and that each agent acts and believes and evaluates in accordance with the same norm he is committed to. Even keeping these ridiculous oversimplifications for the moment, we can do a good deal to accommodate normative debate. For normative debate typically proceeds by citing facts: if the prosecutor is trying to convince the jury that the defendant in a shooting wasn’t justified in believing that his life was in danger, she will try to get the witnesses to convince the jury of straightforward facts about the situation to support this; that is, the idea is to elicit facts that together with the various epistemic norms  $n_1, \dots, n_{12}$  employed by the twelve jurors lead to the conclusion that the defendant wasn’t justified. (Slightly more accurately, the idea is to alter the jurors’ probability functions  $P_1, \dots, P_{12}$  in such a way that  $P_1^*_{n_1}, \dots, P_{12}^*_{n_{12}}$  all assign to the normative claim a sufficiently high value to produce a vote to convict.) Different people differ in their epistemic norms, but there is sufficient similarity in their norms for this process of altering others’ evaluations in desired ways by altering their beliefs to be fairly effective.

The point would be little altered if we allowed the agents’ norms to be imprecise, thereby

---

<sup>24</sup> Indeed, not just inconsistencies: also *conditional* inconsistencies, i.e. giving inconsistent recommendations *on specific factual assumptions*.

*slightly* reducing the level of unrealisticness. An imprecise norm can be viewed to first approximation as a set  $N$  of precise ones, so any probability function  $P$  over non-evaluative claims generates a non-empty set of probability functions over evaluative claims,  $\{P^*_n \mid n \in N\}$ . The remarks on rational debate in the previous paragraph clearly extend: we are in effect just allowing that a single juror might be in a state of mind corresponding to the collective state of mind of the jury in the above.

Before moving to a more realistic picture, I should remark that we might augment the “set of precise norms” picture by imagining that the agent *weights* the norms in his set, by a measure  $\nu$  over norms, where the measure of the entire set of norms is 1. Relative to any such measure, a probability function  $P$  on non-evaluative beliefs generates a unique probability assignment  $P_\nu^*$  to evaluative beliefs: the weighted average of the  $\mu(A/n)$ , with weights given by  $\nu$ . (Formally,  $\int \mu(A/n) d\nu$ .) This would allow for standard probabilistic reasoning to extend to the full normative language, just as the restriction to a single precise norm would. Still, the probabilities involved in the reasoning decompose into two components, the probability measure  $\mu$  over worlds and the one  $\nu$  over norms.<sup>25</sup>

**10. Assessor relativism.** The previous section is related to the distinction between contextual relativism and assessor relativism, as I conceive it, which is at bottom a pragmatic distinction.

I’ll begin by describing MacFarlane’s distinction as I understand it,<sup>26</sup> so that I can relate mine to his. MacFarlane makes his distinction in the context of a semantic framework in which there is an important distinction between *indexical contextualism*, on which the same sentence expresses different propositions in different contexts, and *non-indexical contextualism*, where the proposition itself is evaluated as true only relative to a certain kind of contextual parameter: a “user parameter”. (I expressed some doubt about the distinction between the two types of contextualism in note 4, but will go along with it.) One example of a user parameter (he says) is the possible worlds parameter: if I say “Snow is white” and an English-speaking inhabitant of

---

<sup>25</sup> It would be somewhat more realistic to generalize from ascribing to each agent single measures  $\mu$  and  $\nu$  (on the worlds and precise norms respectively) to ascribing a pair of (non-empty) sets of  $\mu$ -measures and  $\nu$ -measures, or a (non-empty) set of pairs of  $\mu$ - and  $\nu$ -measures. But psychological realism is better pursued by the means suggested in Section 11.

<sup>26</sup> Here I have been greatly helped by both Stewart Cohen and Gary Ebbs.

another possible world says “Snow is not white”, we are evaluating the same proposition relative to our different worlds. Another case that he regards as *natural* to treat this way is the time-parameter in present-tense utterances: rather than supposing that ‘It’s now raining’ expresses different propositions at different times of utterance, we treat it as expressing the same propositions at all times, but we evaluate the truth of the proposition only relative to times. A frame-of-reference parameter for simultaneity could be treated this way too. The propositions in all these cases are “incomplete” (MacFarlane forthcoming, Section 6) in that they are true relative to some parameters of evaluation but not others.

But MacFarlane thinks that propositions can also be incomplete with regard to another kind of parameter, an “assessor parameter”. Assessor parameters function in a very different way. MacFarlane’s terminology might suggest that the distinction is drawn in terms of the different values that the parameter takes on: e.g., that the distinction between propositional truth being relative to a user-parameter for norms and its being relative to an assessor-parameter for norms is that in one case we have relativity to the user’s norms and in the other case relativity to an assessor’s norms. But that would make no sense: then asking whether we had user-relativity or assessor-relativity would be like asking whether simultaneity of events A and B is relative to event A, or to event B, or to the observer. When you have a relative predicate, you can relativize to anything you like: the point of relativism is that none is privileged. The distinction between types of relativity, then, must be not in terms of what the values of the parameter are but in terms of how the parameters function.<sup>27</sup>

And part of that difference seems clear. Even if “It is raining now” expresses a proposition that is “incomplete” with respect to world and time, there is still a non-relative standard of objective correctness for the assertion of or belief in such a proposition on any given

---

<sup>27</sup> Since the distinction is pragmatic rather than semantic, I’m not sure that there’s much point in thinking in terms of two kinds of parameters: we can instead think of one kind of parameter that can be given two different pragmatic uses. Indeed, the latter is probably preferable, for two reasons. First, the two-kinds-of-parameter picture encourages the view that the pragmatic distinction is quite hard-and-fast, and that is probably unrealistic. Second, for anyone who isn’t a possible-worlds realist, there is no real distinction between the two uses of a world-parameter: in MacFarlane’s terminology, the “world of assessment” and the “world of use” must be the same (there’s only the actual world); or in the alternative terminology suggested immediately below, truth and objective correctness coincide. So in the case of worlds (possible worlds realism aside), the two-kinds-of-parameter view makes for a distinction without a difference.

occasion: the assertion or belief is (*objectively*) correct if the proposition is true *with respect to the world in which it is located and the intended time*. (I say *objectively* correct, to distinguish it from a notion of correctness where an assertion is correct if it reflects the asserter's beliefs.) The same holds with relativity as to frame of reference: if you want to say that the proposition expressed is incomplete with respect to frame of reference, still a simultaneity claim is still evaluated as correct or incorrect on the basis of the frame of reference that the speaker intends. I think that the cases where we evaluate a belief or utterance as correct or incorrect on the basis of the parameter that the speaker intends correspond to what MacFarlane has in mind by user-relativity, or non-indexical contextualism.

Assessor relativity is different. But how? An idea that one might have is that here, the assessor's norms rather than the user's go into non-relative assessments of objective correctness. But this view would be quite mysterious, for reasons I won't discuss. I think MacFarlane agrees, and the last part of his 2005 is an attempt to sketch an alternative. I don't fully understand it; what follows is probably a different view, but I think is not far from the spirit of his.

In assessor relativity as I shall understand it, there is no non-relative notion of objective correctness at all—at least, none that plays any significant role in normative discourse. The only notion of correctness that plays a serious role in normative discourse is truth, and it is norm-sensitive.

How is truth to be understood? On my understanding of 'true', the claim that the proposition  $p$  is true is equivalent to  $p$  itself: if  $p$  has a norm-sensitivity,  $\text{True}(\langle p \rangle)$  has exactly the same norm-sensitivity, and if  $p$  doesn't then  $\text{True}(\langle p \rangle)$  doesn't either. When  $p$  or  $\text{True}(\langle p \rangle)$  is norm-sensitive, then you can de-sensitize, by making the relativity explicit. One way to do this is user-focused: we prefix by 'Relative to the user's norms ...'. Another way to de-sensitize is assessor-focused: we prefix by 'Relative to the assessor's norms ...'. There are also other, and I think more important, ways to de-sensitize: e.g., 'Relative to any set of norms that allows for learning from experience in such and such a way at such and such a rate, ...'. But *assessor-relativity doesn't have anything to do with de-sensitizing in one of these ways rather than the other*. Assessor-relativity doesn't involve de-sensitizing the normative propositions at all. Nor does it involve de-sensitizing claims of the truth or correctness of normative propositions.



Attributions of truth to such normative propositions are norm-sensitive in just the way that the ground level normative claims are; and (what distinguishes the view from standard cases of “non-indexical contextualism” as MacFarlane understands it) no norm-insensitive notion of “objective correctness” plays any substantial role in the assessment of normative utterances. It is part of the pragmatics that the utterer or assessor uses her own norms to decide what to accept, but objective correctness in any sense distinct from norm-sensitive truth simply doesn’t enter into it.<sup>28</sup> (I’ve heard it objected that on this view the parameter for norms functions like a free variable, so that what’s asserted is propositional functions rather than propositions. But these “propositional functions” are evaluated in terms of truth, understood as itself norm-sensitive; and as discussed in section 6, they can be the objects of impure belief. For these reasons it seems reasonable to call them propositions.)

What makes the “incompleteness” far more important for norms than for worlds, times, or frames of reference is that they connect up with the phenomenon of disagreement. In the cases of contextual relativity like ‘It is raining’, people don’t count as disagreeing unless they disagree in their straightforwardly factual beliefs. But this is not so in general: two people can disagree about where to go to dinner, even though there is no relevant factual disagreement between them. So too in the normative case: people who advocate doing different things, or make opposed claims about what they ought to do, count as disagreeing, even if the difference stems not from a difference in their straightforwardly factual beliefs but from a difference in those policies or preferences that generate normative claims. Typically, the disagreement about a specific matter, e.g. how quickly the government ought to withdraw troops from Iraq, will be due both to straightforwardly factual differences and to rather basic normative policies. Disagreement concerns opposed *impure beliefs*, which are resultants of both factual beliefs and norms.

---

<sup>28</sup> There is an ambiguity in ‘non-indexical contextualism’. The usual official explanation is semantic, but this semantics can have two very different pragmatics associated with it, and I think that MacFarlane and most others who use the term associate with it the pragmatics on which normative utterances are taken to be objectively correct iff they are true relative to the utterer’s norms. As long as one recognizes the alternative pragmatics, which employs no such notion of objective correctness but only a norm-sensitive truth predicate, then it makes no difference whether one calls this a variety of non-indexical contextualism or reserves the term ‘non-indexical contextualism’ for views with the usual pragmatics.

As a matter of psychological fact our norms and our beliefs are seamlessly integrated. To what extent the disagreement is based on “straightforward facts” and to what extent on “basic norms” is extremely hard to determine in practice. (So a philosopher who for some reason prefers to reserve the term ‘genuine disagreement’ for straightforwardly factual disagreement would be advocating a notion of disagreement that would be hard to employ in practice, because it would be extremely hard to determine when people disagree in the proposed sense.) This is part of what underlies the naturalness of talk of impure belief: we can say in general that disagreement consists of having opposed impure beliefs, without differentiating the contributions to impure belief made by pure (i.e. straightforwardly factual) beliefs on the one hand and by norms on the other.

The pragmatic difference between “assessor relativity” (to appropriate MacFarlane’s word for the distinction as I’ve drawn it) and ordinary contextual relativity is of fundamental importance. Because of it, the distinction between implicit relativity and explicit relativity is of vastly more significance in the normative case (where it is assessor relativity that is in question) than in the contextual relativity examples. In the case of contextual relativity, two people disagree in their utterances of an implicitly relativized sentence such as ‘It’s raining’ if and only if they would disagree had they explicitly relativized to the locations that they intended. So it really makes little difference whether they make the relativity explicit: in leaving it implicit they don’t in any important sense say anything different than they would have had they made it explicit. This is not so in the case of assessor relativity, for here explicit relativization removes the norm-sensitivity. If Jones and Smith utter

J: We ought to withdraw our troops within a month

and

S: We ought not to withdraw our troops within a month,

they are clearly disagreeing, due to some combination of disagreement about straightforward facts and disagreement in fundamental policies. They have opposed impure beliefs. But if they explicitly relativize by uttering

J\*: We ought to withdraw our troops within a month relative to Policy  $n_J$

and

S\*: We ought not to withdraw our troops within a month relative to Policy  $n_S$ ,

where  $n_J$  and  $n_S$  are fundamental policies they advocate, they may not be disagreeing: for instance, if they agree on the straightforward facts, then Jones will agree with  $S^*$  and Smith will agree with  $J^*$ . In short,  $S^*$  and  $J^*$  are straightforwardly factual claims: the sensitivity to norms has been lost by the explicit relativization. And sensitivity to norms is a large part of what normative disagreement consists in.

I've made the point with reference to a particular method for specifically relativizing, but there are many other methods for doing so, and the point extends to other such methods. For instance, suppose Jones and Smith value norms with certain properties, say  $\psi_J$  and  $\psi_S$ . Then they may be tempted to explicitly relativize not by  $J^*$  and  $S^*$ , but as follows:

$J^{**}$ : We ought to withdraw our troops within a month relative to any policy with property  $\psi_J$

and

$S^{**}$ : We ought not to withdraw our troops within a month relative to any policy with property  $\psi_S$ .

It may be the case that  $\psi_J$  and  $\psi_S$  are the same, in which case the disagreement between Jones and Smith would be entirely factual and would be preserved if they were to utter  $J^{**}$  and  $S^{**}$  instead of  $J$  and  $S$ . But it may be that they are not the same. In that case one of  $J^{**}$  and  $S^{**}$  could still be false, so that Jones's and Smith's disagreement could still be a matter of one having false purely factual beliefs. But it could be the case that Jones and Smith know all the relevant facts, and that  $J^{**}$  and  $S^{**}$  are both true. So it is still the case that in relativizing to  $J^{**}$  and  $S^{**}$ , the disagreement can be lost.

I've said that in the case of contextual relativity it makes little difference whether speakers make the relativity explicit: in leaving it implicit they don't in any important sense say anything different than they would have had they made it explicit. It may be asked whether in the case of assessor relativity they say anything different when they explicitly relativize than when they don't. And it may be thought that I'm in trouble whatever answer I give:

If you say that they say the same thing when the relativization is implicit as when it is explicit, then you can't sensibly say that they can disagree in the first case when they agree in the second. If on the other hand you say that they say something different when

the relativization is implicit than when it is explicit, then the truth conditions must differ in the two cases, and so the relativism doesn't really capture the truth-conditions being assigned in the normative ("assessor-relative") case.

But this objection turns on terminology ("says the same thing as") that must be treated carefully.

My primary response to the objection is to the second ("If on the other hand ...") horn. Consider the *best* case for the claim that J "says the same thing as" J\* or J\*\*: the case where Jones consistently advocates and employs norm  $n_J$ , or consistently advocates employing whatever norm satisfies property  $\psi_J$ . Even then, it would strike me as highly misleading to say that Jones "says the same thing" when uttering J as when uttering J\* or J\*\*. J says something quite different from either J\* or J\*\*, since J can be used to express normative disagreement with Smith in the way that J\* and J\*\* can't. To put it another way, J says different things than J\* and J\*\* because one can impurely disbelieve J without impurely disbelieving J\* or J\*\*. Moreover, there's a perfectly clear sense in which J differs in truth conditions from both J\* and J\*\*. After all, J has only norm-sensitive truth conditions, whereas each of the two explicitly relativized sentences has norm-insensitive ones; this is a semantic reflection of the fact that only the explicitly relativized sentences can be objects of *pure* belief. And relative to some norms (e.g. those that both are different from  $n_J$  and don't have property  $\psi_J$ ), J differs in truth conditions from both J\* and J\*\*: that is why it "says something different" from J\* and from J\*\*. The norm-sensitive truth conditions thus *do* accord with the assessor-relativism, contrary to what is asserted at the end of the second horn of the dilemma.<sup>29</sup> The view that there is a discrepancy between the ascribed truth conditions and the assessor relativism comes from trying to read norm-insensitive truth conditions into the relativism, but that is inappropriate—especially in the case of assessor-relativism, where the distinction between implicitly relativized statements and their explicitly relativized counterparts is so crucial.

---

<sup>29</sup> In saying that the assessor-relativism accords with the norm-relative truth conditions, I don't mean to imply that one could explain the difference between assessor-relativism and contextual relativism in terms of truth conditions. That is certainly false. It is central to the transparent notion of truth that I've been assuming that meaning can't be explained in terms of truth conditions; meaning "comes first", and truth conditions reflect it. The distinction between assessor-relativity and contextual-relativity is, as I've said, a pragmatic distinction, having to do with their different roles as regards the social phenomenon of disagreement.

Is there a sense in which the *straightforward factual content* of J, as uttered by a person who consistently advocates and employs norm  $n_J$  (or consistently advocates employing whatever norm satisfies property  $\psi_J$ ) is no different than that of J\* (or J\*\*)? Perhaps there is, in some cases (though in typical cases where one both advocates particular normative views and particular features of norms, it would be difficult to decide which particular style of explicit relativization is appropriate). If so, and if by ‘truth conditions’ one means *straightforward factual content* in this sense, then J does indeed have the same truth conditions as one of J\* and J\*\*; and if “saying the same thing” means having the same straightforward factual content in this sense then J does indeed say the same thing as the chosen explicitly relativized sentence. But this seems to me a misleading way to use the term ‘truth conditions’, and an even more misleading way to use the term ‘say the same thing’. It is even a misleading way to use ‘straightforward factual content’: it would be better to deny that normative utterances have straightforward factual content, or to say that they do only relative to norms. (One reason for adopting one of these last two options is that most people don’t consistently employ precise norms or impose precise requirements  $\psi$  on them, so that it is hard to see how to assign a “straightforward factual content” of their normative utterances that is not itself norm-sensitive.) Still, if one persists in talking in accordance with the first horn of the dilemma, the answer to it is that normative disagreement just isn’t a function of “what is said” in this straightforward factual sense: normative disagreement isn’t a matter of opposed *pure* beliefs, it’s a matter of opposed *impure* beliefs.

If as I’ve suggested the distinction between assessor relativism and contextual relativism is pragmatic, won’t it admit of borderline cases? I think it will, but that this is quite appropriate. Indeed, many disputes about taste seem to be borderline in much this way. If someone exhibits preferences in food or wine or music that I don’t share, I may just regard his preferences as different: he likes peanuts, I don’t, but there’s nothing we disagree about. But in other cases I do regard us as disagreeing: I certainly disagree with anyone so misguided as to prefer Neil Sedaka to Neil Young. And there are cases in between. (The difference seems to have to do with the extent to which one takes a live-and-let-live attitude to the other’s preferences, which in turn is *in part* due to the impact of his acting on those preferences on one’s own concerns: I’m more inclined to regard myself as disagreeing with someone whose favorite dessert is Rice Krispie

Treats if his plan is to order it for the whole table than if he consumes it among consenting adults in the privacy of his own home. If the person acts on moral preferences different from mine, this will usually affect things I care a lot about, which I take to be a large part of the reason that moral differences are usually disagreements whereas differences in taste often aren't.)

**11. The modification of norms.** The initial picture of normative debate given in Section 9 offered no dynamic for how debate might lead to our changing the norms that guide our epistemic (belief-forming and belief-retaining) behavior. For a sufficiently “low level” norm, there’s no mystery about this: one thing that people can debate is the likely effects of a policy, and for sufficiently “low level” norms it’s easy to see why we would alter them if we thought that they were unlikely to lead to desirable results. For sufficiently “high level” norms, though, (and especially for “highest level norms”, if you believe there to be such a thing), this is problematic.

An initial though somewhat subsidiary point is that it’s doubtful that “high level” norms guiding our epistemic behavior should be viewed as goal-oriented. Indeed, if being goal-oriented is understood literally in terms of our attempting to achieve the goal, the idea seems incoherent: we’d need to form beliefs about how the goal is best achieved, and a regress would ensue. Even on a more liberal construal, another problem is to figure out what the goal might be. For instance, what is the goal in the case of inductive norms? It is often said that the goal is some tradeoff between reliability and power; but these terms (especially ‘reliability’) have little clear meaning in application to rules that can be applied in the wide range of circumstances that candidates for inductive rules can be applied. Part of the problem is that inductive rules are highly “self-correcting”: if applied in circumstances in which they initially behave unreliably, they typically start to behave more reliably. This is so for many bad inductive rules 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”.<sup>30</sup>

I don’t doubt that there are things we can say about what makes one inductive norm

---

<sup>30</sup> There’s a bit of discussion of this in Field 2000, Section 4.

better than another. For instance, some inductive norms (e.g. Carnap's  $\lambda$ -continuum with very large  $\lambda$ ) are bad in that they "self-correct" too slowly. Others (e.g. Carnap's  $\lambda$ -continuum with very small  $\lambda$ ) are bad in that they "self-correct" too *quickly*: they make predictions so close to observed frequencies, or to recent observed frequencies, that they are too easily influenced by statistical flukes. (In addition, *all* of Carnap's methods are bad in that they are very limited in the kind of evidence that they can take into account.) But saying that our norms are to be evaluated by these and a variety of other "truth-oriented" criteria is much weaker than saying that we adhere to them because we think they are the best means of achieving some tradeoff among those truth-oriented criteria.

Still, the idea that we debate norms by debating whether they are likely to lead to desirable results—in particular, truth-oriented results of various sorts—does not require the stronger construal in terms of a goal: it is enough that there be a variety of desirable features (presumably having some sort of relation to achieving truth and avoiding falsity) with respect to which we evaluate them. But this brings us to the main point: for sufficiently high level norms, it seems problematic how there can be rational debates about this. Won't each method automatically predict that it will do better than the alternatives? (The *problem of immodesty*.) If so, each of the debating parties would seem to be unpersuadable.

The problem of immodesty is probably a bit overstated. Attempts to prove that every method declares itself better than all others tend to adopt simplistic 1-dimensional criteria of what is good in a method. Moreover, even given the criteria for what is good, the arguments depend on controversial measures of closeness to the good: as shown in Gibbard 2008, alternative measures of closeness do not yield the result that all methods declare themselves best. Still, it's hard to take any comfort in this: the alternative measures of closeness, on which not every method declares itself best, seems to be ones on which the only methods that do declare themselves best are exceptionally bad ones! (See the Gibbard paper for details.) There is no reason whatever to think that there is a way to spell out a criterion of betterness (i.e. of goodness and of closeness to the good) on which most methods that ought to come out bad will *declare themselves* to be bad but on which methods that ought to come out good will declare themselves good. So while the problem of immodesty is overstated, there does seem to be a deep underlying

problem with the idea that rational debate between consistent advocates of alternative “highest level norms” would lead to the better method winning out over the worse.

There seems to be in addition another problem, the *problem of modesty*. Suppose that by making factual judgements in accordance with a certain epistemic norm  $N$ , I arrive at the view that norm  $N^*$  is better than  $N$  according to a certain criterion of betterness. (Assume that application of  $N^*$  and  $N$  to our observations leads to conflicting judgements on at least some factual claims.) And suppose that as a result of this, I switch to norm  $N^*$ , and start making factual judgements in accordance with  $N^*$  rather than  $N$ . It seems that if this is a rational switch, then  $N$  can't have been the highest level norm guiding my behavior after all: if following  $N$  was really my highest level policy, then I would have stuck with it, despite the fact that it came out worse on a certain criterion of betterness.<sup>31</sup> So at least for those who think that there is a highest level epistemic norm that guides our behavior, it would seem that that norm can't be subject to change by rational debate.

That's the argument, but it can't be right as it stands. Suppose that  $N$  was indeed my “highest level policy”. Why does that imply that if, by following  $N$ , I came to the conclusion that  $N$  is worse on a certain criterion of betterness, I would have stuck to  $N$ ? What's true is that if I then started following the policy I thought was better, I would no longer be following  $N$ . So *if my abandoning  $N$  was norm-driven*, then  $N$  wasn't my highest level norm. But why suppose it had to be norm-driven?

It may be said that what's at issue isn't arbitrary change of norm, but *rational* change of norm, so that the argument can be repaired by altering the claim to: “if following  $N$  was really my highest level policy, then *barring a non-rational change of highest-level norm* I would have stuck with it, despite the fact that it came out worse on a certain criterion of betterness.” But then the argument is question-begging: it presupposes that any change of highest-level norm is

---

<sup>31</sup> Norms can include preferences as well as policies, but these are different kinds of norms. The assumption that  $N$  is the highest-level policy-norm guiding  $S$ 's behavior does not preclude that  $S$  evaluates it, but it does preclude that  $S$  has a norm to change  $N$  in light of such an evaluation. It is even conceivable to retain a policy-norm despite a *highest level* evaluation that one shouldn't: consider the theoretical skeptics who (apparently) have norms of evaluation that disvalue the policy-norms that they employ. This is highly undesirable, and casts serious doubt on the utility of the norm of evaluation, but it is not incoherent.



non-rational.

One thing that does seem right in the problem of modesty is that no policy, highest level or not, can dictate its own revision: so if it can be rational to change a highest level norm, this can't be explained by saying that it was what the norm told us to do. But it isn't obvious why we should have to explain it in that way.

What is at issue here is a change in which norm *guides us*. The most natural idea of how such change might come about is for us to first come to *advocate* (i.e. *be committed to*) our using a different norm, and to then train ourselves to act or reason in terms of the norm we have come to advocate instead of the previous norm. The rational debate will be first and foremost a debate about what norm to advocate (be committed to). If as a result of such debate we come to advocate a norm that we know differs from the one that has guided us, it is natural to try to bring about a change in how we are guided. We may or may not be able to do this; and if we succeed in doing it, there may be an issue about whether in making the change in norm we are acting rationally. I'll have more to say about that last issue near the end of this section. For now I just want to stress that the issue of whether the change in the norm that guides us is rational is a distinct issue from the question of rational debate; if there is rational debate, it is over the norms to advocate.

A. On the question of *rational debate* about which norms to advocate, there still seems to be something akin to the problem of immodesty: if we advocate a certain epistemic norm *N* as a "highest level norm", it is not clear that we can give a standard for evaluation such that

(i) consistently following *N* would lead to the evaluation that *N* is superior to its rivals,

and yet

(ii) consistently following a typical rival would *not* lead to the evaluation that that rival is superior to *N*.

And this may seem to preclude rational debate between those who advocate different norms.

I think that this last conclusion is incorrect. Among other things, it turns on a far too idealized picture of how rational debate proceeds. To get a model for how debate about any deeply entrenched belief or norm proceeds, we should abandon the assumption implicitly made

in Section 9, that agents are logically omniscient. Actual agents are far short of this: they do not see all the consequences of their beliefs, policies and preferences, and this leads not only to many failures of logical closure but also to many unrecognized inconsistencies. In the earlier discussion, I talked of the cognitive state of an agent as represented by a “pure component” that consists of a measure on a set of worlds or a set of such measures, together with a normative component that consists of a precise norm, or a set of precise norms, or a measure on the precise norms, or a set of such measures, or whatever. This is an acceptable idealization in circumstances where inconsistencies play no role; but in contexts where inconsistencies do play a role, it blinds us to how rational debate proceeds.

It seems clear in rough outline how to construct a better picture. In the case of pure belief, we can suppose that at any time, an agent  $X$  has a certain body of *core doxastic attitudes* toward non-evaluative claims. The exact form of these core doxastic attitudes depends on whether one adopts a picture of the agent as having something like degrees of belief (as opposed to absolute beliefs plus “spheres of implausibility” to handle belief revision). Even on the degree of belief picture, the core attitudes needn’t have the simple form of attributing specific degrees of belief to specific claims, they can be far more complicated: e.g. degree of conditional belief (like one’s degree of belief that heads will result *given that the coin is flipped*), or upper and lower bounds on degree of conditional belief, or comparative degree of conditional belief (“ $A$  is more likely given  $C$  than  $B$  is given  $D$ ”), and so forth. This set of attitudes will not be deductively closed (or closed under probabilistic consequence): one has no degree of belief in very complicated logical truths that one hasn’t explicitly contemplated. And it needn’t be (and in practice won’t be) consistent or probabilistically coherent: e.g. an agent might well have a degree of belief far less than 1, even 0, in some complicated logical truth.<sup>32</sup> One might invent methods for assigning to such an agent a set of probability functions to represent the pure doxastic state (roughly speaking, taking the agent’s pure doxastic state to be the set of probability functions that satisfy *sufficiently many* of the agent’s core doxastic attitudes), but satisfactorily working out the details of this would be complicated, and the result might do more to obscure important

---

<sup>32</sup> Such as the conditional whose consequent is that there are continuous functions mapping the unit interval onto the unit square and whose antecedent consists of the axioms of set theory plus the set-theoretic definitions of the notions appearing in the consequent.

features of rational debate than to illuminate them.

We can extend this to impure doxastic attitudes, either directly (attributing to the agent a certain body of *core impure doxastic attitudes*, toward *non-evaluative and evaluative claims together*); or indirectly, by attributing to the agent a certain body of core policy-commitments, core preferential commitments, etc., which together with the pure beliefs will generate the impure attitudes. Either way, the set of attitudes won't be deductively closed, or consistent. Indeed, the presence of the pure beliefs and the policies together vastly increases the routes to inconsistency from those present in either one alone, for a set of policies may be consistently carried out only in worlds very different from how one knows the actual world to be.

Again, one can invent methods for assigning to an agent subjective measures  $\mu$  and  $\nu$  on the worlds and norms respectively, or rather, sets of pairs of such measures: the idea, as before, would be to look at sets of pairs  $\langle \mu, \nu \rangle$  that satisfy sufficiently large subsets of the agent's impure core. But again, I think that an understanding of the dynamics of doxastic states is better achieved without this. The reason for this—as in the case of pure doxastic attitudes—is that the measures are an epiphenomenon: the real work goes on at the level of the core attitudes. If the core attitudes were consistent and evolved in accord with an extremely demanding picture of “idealized rationality”, the measures (or sets of pairs of measures) would evolve in a smooth way that could be described without mention of the underlying core; but since the core attitudes aren't even consistent, the evolution of the measures won't be characterizable without reference to the underlying core.

I'm in no position to give anything like a complete account of how these inconsistent cores rationally evolve, but even without that, I can say something about how this picture opens up the possibilities for rational normative debate. The point is obvious: in normative debate one (consciously or unconsciously) exploits inconsistencies and other tensions in the other person's normative views, or between their normative commitments and the norms they employ in acting/believing or in evaluating. (Something analogous happens in non-normative debate: we exploit inconsistencies and other tensions in non-normative views.) By a tension in one's views, I mean an uncomfortable commitment: the views may be consistent, but only given a further claim which the agent might be uncomfortable in accepting or be brought not to accept. In

debate we typically argue, from things the person explicitly accepts and other things he can be easily brought to accept, to something *A* that is in conflict with what he accepts. The person may resist the argument, by questioning some of the claims used in it (even ones he previously accepted); but a good arguer is likely to find other ways to argue for *A* from things he accepts, and with enough such argument the person is likely to be persuaded to alter his views and accept *A*. If the person had debated someone else, he might well have been led to resolve the inconsistency in his views by keeping *not-A* and altering some of the related views. This is how, for instance, people faced with the semantic paradoxes may end up with different choices about whether to keep their logic and revise their basic rules for truth, or to keep their basic rules for truth and alter their logic.<sup>33</sup> This process of revision can lead to fundamental change in both pure and impure core attitudes. And in the latter case, if the overall change is important enough, it will constitute a change in the norm advocated (since which norm one advocates is determined by which norm provides the best fit for one's doubtless-inconsistent set of impure doxastic attitudes). And if the person succeeds in adopting the policy he advocates, it will constitute a change in the norm that guides his behavior.

It may be thought that in focusing on inconsistency within norms I am illicitly thinking of norms as explicitly represented in us. The norms in question are high-level policies of some sort; or if you like, *rules*. But as is frequently pointed out, we can't suppose that all of the rules that an agent employs are explicitly represented in their heads: that would lead to a familiar regress, in that we would then need policies or rules for processing the internal representations.<sup>34</sup> Some of the policies or rules that we describe a person as employing are merely implicit in the person's practice. But this means that they result from a kind of idealization in the person's practice. And

---

<sup>33</sup> I take this example to be normative: the issue is about *the rules by which to reason*. (See Field 2009.)

<sup>34</sup> Well, we could stipulate that the processing of the representations shouldn't count as literally rule-governed since it is not represented; but this stipulation seems insufficiently motivated. It seems reasonable to regard someone as "following" a rule when (i) the person's behavior by and large accords with the rule, and there is reason to expect that this would continue under a decent range of other circumstances; and (ii) the person tends to positively assess behavior that accords with the rule and to negatively assess behavior that violates the rule. (In the case of epistemic rules, the "behavior" is of course the formation, retention, or revision of beliefs.) This is vague, which reinforces the point next to be made, about the considerable indeterminacy involved in ascribing epistemic or other rules to a person.

it is at least somewhat natural to suppose that the process of idealization imposes consistency, so that we can't reasonably suppose that the high-level rules governing a person's epistemic behavior are inconsistent.

While I'm somewhat skeptical of the last step, there's no reason to take a stand on it: the argument is really quite beside the point. For the main point under discussion in the paragraphs prior to the previous one was rational debate about what norms *to advocate*, and those will be explicitly represented: it's hard not to explicitly represent what one debates about. There is no doubt that one can *advocate* norms that are in fact inconsistent, even if one can't be guided by them. (I did at one point in the argument mention inconsistencies between the norms one advocates and the norms that guide one, but did not suppose any inconsistency within the latter.)

**B.** Let us now move from the question of rational *debate* about the norms advocated to the question of rational *change*, both in norm advocated and in norm employed. Under what conditions will change produced by rational debate be rational? There can be rational debate in which a person with a faulty position persuades one with a better position to abandon it in favor of the faulty one; this can happen when both parties are debating in good faith, and it can happen both in normative and non-normative debate. There are cases of this sort where I'd have considerable inclination to say that the change from the better norm to the faulty norm is rational, despite it being a change for the worse: these are primarily cases where the argument is extraordinarily compelling, and the resultant position not a whole lot worse than what it replaced. There are other cases where I'd have considerable inclination to say that the change from the better norm to the worse was irrational, despite it being on the basis of an argument that we can't blame the convinced party for not seeing through: these are primarily cases where the faulty norm is *much* worse than what it replaced, and where the convinced party may have had some grounds for suspicion. There are cases where I feel conflicting inclinations: for instance, someone convinced by a highly persuasive but faulty argument for adopting a statistical procedure that in fact is deeply flawed. Do we need a theory of when such changes are rational and when they aren't, that will decide such cases?

Well, I guess on the dipstick model we do need such a theory: we need to know how much fluid is produced by a persuasive but faulty argument, and how much of this is then

pumped out by any “rational intuitions” against the conclusion, so that we can determine the overall level. But absent such a model, I think the thing to say is that what we are doing in epistemology is giving *multi-faceted* evaluations.<sup>35</sup> Suppose I see that the statistical procedure that Jones used to employ, before he was convinced by Smith’s incorrect argument to change it, was a good one, and suppose that I know how Smith’s argument went wrong; still, I see the apparent power of Smith’s argument to anyone not immersed in subtle issues in the philosophy of statistical inference. I positively evaluate Jones’s intellectual honesty in following out reasoning that seems persuasive, and being willing to revise his work in light of this, especially when the fallacies that led him to do this are subtle. At the same time, the costs of employing Smith’s faulty statistical procedure may be fairly high, so I think there is also something quite negative about Jones’s conversion. Why think that more needs to be said? Why think there needs to be a single standard of reasonableness, and that these two factors need to be weighed against each other? From my point of view, Jones is in an unsatisfactory overall credal state, but weighing each belief in the state on its own and on a single scale is pointless.

These last remarks apply both to change in the norms one advocates (is committed to) and also to change in the norms that guide one. I remarked above that the usual way to change the norms that guide one is to first come to advocate new ones and to then train ourselves to act or reason in terms of them. When is the last step rational? This I take to not be a straightforwardly factual question, but a question of evaluation. There are cases where I’d be more inclined to evaluate the *advocacy* of a change in norms as rational than to evaluate the *employment* of the change of norms as rational: cases where the person is seduced by persuasive arguments for skepticism may fit in here. There are also cases where a person may change the norms that guide him not as a result of rational argument for changing the norms he advocates, but in some other way: perhaps by a change of the norms that guide him unaccompanied by a change in the norms he advocates, or perhaps as a result of change in the norms advocated

---

<sup>35</sup> I don’t claim that anyone who believes in metaphysical justification (“justificatory fluid”) needs to think that there is a single overall level of fluid that matters, as opposed to levels at seventeen different places (or the amounts of seventeen different fluids). Still, I think that the belief in metaphysical justification does tend to go with belief in a single standard, or at least a most privileged one, and so I have taken the “dipstick model” to include this additional commitment.

because of an a-rational change in preference. For instance, we might come to prefer inductive policies that self-correct more slowly, or that are more cautious about accepting generalizations; and such changes of preference might lead us to “retrain ourselves” to follow inductive policies more in accord with these new preferences. (As remarked above, the inductive policies are not literally goal-driven, so there is no obvious argument that such a change in inductive policy would be guided by a more fundamental norm.) In these cases, I’d probably count the change as rational if the new guiding norm seemed to me far superior to the old one. But in my view there is really no issue worth debate in these cases: it’s a matter of evaluation, not of metaphysical fact.

**12. “Highest level” norms?** I have argued that change we deem rational needn’t be change that was produced by following a norm of rationality. I would like to note that this undercuts one argument for belief in “highest level norms” guiding our behavior. As background, recall that when we ascribe a high-level epistemic norm to someone we are giving an idealized description of how the person forms and alters beliefs. This idealized description needn’t be unique: one reason for this is that the doxastic core is inconsistent, and there may be more than one equally good way of getting an “approximate fit” by a consistent norm. Another reason, probably more important, is that we can idealize to different degrees: for instance, some idealizations take more account of memory limitations or computational limitations than do others. So there are multiple good candidates for the best idealization of our epistemic practice.

This does not itself go against the idea of a highest level norm guiding our behavior; indeed, it somewhat removes the *prima facie* implausibility of that. Since in attributing norms one is idealizing, the issue of a highest level norm is the issue of whether *a good idealization will postulate* a highest level norm; this is compatible with different good idealizations postulating different highest level norms (whether of the same or different “degrees of idealization”). And I think that much of the *prima facie* implausibility in the idea of a “highest epistemic norm” governing our behavior is removed by the idea that there is no uniquely best candidate for what this norm is.

Still, the question remains, why should we believe that the best idealizations will posit a highest level norm? The answer that I gave in the Appendix to my 2000 paper [with some minor

alteration to fit the terminology of the present paper] was that

...the alternative is an idealization that postulates multiple norms, each assessable using the others. But there is an obvious weakness in an idealization of the latter sort: it is completely uninformative about what the agent does when the norms conflict. There is in fact some process that the agent will use to deal with such conflicts. Because this conflict-breaking process is such an important part of how the agent operates, it is natural to consider it part of a norm that the agent is following. If so, it would seem to be included in a basic or highest-level norm, with the “multiple norms” really just default norms that operate only when they don’t come into conflict with other default norms. Of course, the process of resolving conflicts provided by this basic norm needn’t be deterministic; and as stressed before, there need be no uniquely best candidate for what the higher norm that governs conflict-resolution is. But what seems to be the case is that idealizations that posit a basic norm are more informative than those that don’t.

In retrospect, it’s hard for me to see the force in this, for two reasons.

One reason is that I’ve granted that there are different degrees of idealization (and I discussed this at greater length in that paper). So why shouldn’t the process that decides the conflict within a norm be excluded from that norm, but included only in a norm of lower degree of idealization (i.e. one that takes more account of computational structure)?

The second reason is that the idea that the conflict-breaking process ought to be considered part of a norm that the agent is following is suspect (especially in the case where the process is non-deterministic). I think that what must have led me to endorse the idea was the thought that the resolution of such conflicts is intuitively rational, and that therefore we ought to include the process that led to it in the agent’s norms. But if that was the thought, then the argument was a result of not taking my own evaluativism seriously enough: judgements of the rationality of a piece of epistemic behavior are evaluations that take into account lots of factors other than the descriptive factors that go into an account (idealized or otherwise) of how that behavior was produced.<sup>36</sup>

---

<sup>36</sup> I don’t mean to suggest that it is *only* the evaluator who could see the argument as erroneous on this ground.



**13. The impact of evaluativism on “first order” epistemology.** Part of the ground level import of evaluativism is that the function of morality, epistemology etc. is to give advice, to oneself and to others. It must be advice of the right sort: moral advice, epistemological advice, etc., as the case may be. (For instance, if one is considering whether to perform a certain action, which has obvious practical advantages but one has moral qualms about, then if given advice about the advantages to one’s career one will say “Yeah, yeah, I know all that, but what I wanted was advice of a different sort.”) I suspect that most people, even most moral realist philosophers, recognize that the primary function of moral judgement is to give advice of the appropriate sort, not to discuss the presence of supernatural properties (“justificatory fluid of the moral sort”). It isn’t *just* a matter of giving advice, since we make moral claims about people we’re in no position to advise, e.g. because they would have no interest in our advice, or because they’re dead. But I think it’s reasonable to suppose that the function of this is largely derivative: often we are honing our skills at giving advice (to ourselves and to others), by thinking about what we would recommend in those circumstances; or we may be helping others to hone their skills.

I suspect that most non-philosophers probably also treat epistemological discussion in this way, as a matter of advising on what to believe. Many philosophers do too, a lot of the time, but I think that many are sometimes caught up in the supernaturalist picture. Consider skepticism about induction. Hume points out that we can’t give a non-circular argument for induction, and then while in his study seems to worry about whether this makes the practice of inducing irrational. Once outside of his study he realizes that the worry is absurd: he can’t seriously advise himself or anyone else to have no beliefs about the future or about unobserved portions of the present and past. But if you look at epistemology correctly, there is no need to divorce what happens in the study from what happens outside. If (in your study) you ask “Where does the justification come from? What is its source?” you’re asking the wrong question: you’re thinking of justification as like well water. The only thing to ask is whether and how to form beliefs about the future and unobserved portions of the presence and past; any advice to be so falsehood-adverse as to stop forming beliefs is obviously too silly to consider. (The grounds against such extreme falsehood-aversion are doubtless partly practical, but any proposal to divorce epistemological advice so thoroughly from practical advice would make epistemological advice useless. We can properly distinguish epistemological advice from the narrowly pragmatic

advice we might give someone about, e.g., what to believe to improve his self-image, but this doesn't mean that there's a complete divorce between the epistemological and the pragmatic.) It seems to me that Hume himself recognized that debating the rationality of induction would be absurd; those, like Wesley Salmon, who haven't are in the grip of the justificatory fluid picture.

I certainly don't mean to suggest that all the problems about justifying one inductive method over others go away once one abandons the excessively metaphysical picture of justification. My point is only that that picture stands in the way of thinking sensibly about the problems. For on that picture, it is hard to resist the conclusion that our inductive method has no more of a "justificatory source" than does any other "self-justifying" inductive method (e.g. a counter-inductive one); and since according to that picture one is only reasonable in believing things when there is a metaphysical justification for so doing, it is hard to resist the conclusion that our inductive conclusions are no better than the conclusions of a counterinductivist. Admittedly, few advocates of a "justificatory fluid" view would willingly accept this conclusion; they want to end up recommending methods that it is sensible to recommend, and since they think they can only recommend methods that squirt justificatory fluid, they try to come up with views of justificatory fluid that accord with desirable methods. But if this is possible at all, it requires totally *ad hoc* assumptions about justificatory fluid with no metaphysical naturalness, whose only recommendation is that it makes the alleged metaphysical justification coincide extensionally with what a reasonable person would recommend.

For a second example, consider debates about whether logic is justifiable a priori (in the sense of (i) being legitimately usable without empirical evidence and in addition (ii) not being subject to empirical defeat). From an evaluativist perspective, the only question here is whether it is a good idea to reason logically in these ways; the alternatives are (E-i) to refrain from reasoning logically until one has gathered empirical evidence that doing so has desirable properties, and (E-ii) to let empirical considerations count against logical reasoning under certain conditions. (E-i) seems thoroughly unattractive: without reasoning logically one could never evaluate evidence for so doing. Admittedly, (E-i) was worded so as not to require that we *evaluate* the evidence for the logic prior to using the logic, but only that we *gather* it; but what's the point of gathering evidence if we don't evaluate it? (E-ii) is perhaps slightly less

unattractive, but it is hard to spell out in any detail a scenario on which there would be any motivation to revise logic on empirical grounds.<sup>37</sup> One can read the empirical revisability claim in such a way as not to require a serious story, or even to require that it is *genuinely possible* for there to be a story that would stand up to any scrutiny: one can read it as simply requiring that we can't now rule out with absolute certainty that there might someday be such a story. On that weak reading, the empirical revisability claim is hard to dispute. I'd take that as reason to think that we can't rule out with absolute certainty that logic is empirical; if you want to say that it's enough to show that logic *is* empirical, you can, though it strikes me as a perverse way to talk. But in any case, it seems to me that there is no real issue about whether logic is a priori justifiable, other than these decisions about whether to use it without empirical evidence and whether empirical evidence could undercut its use.

For a third example, consider debates about a particular set of logical laws, not on the basis of empirical considerations but on the basis of conceptual considerations such as vagueness or the semantic paradoxes. Suppose that one of the laws in question is excluded middle. If one thinks of the issue as whether belief in excluded middle can have some kind of "metaphysical justification", I think the position of the classical logician would be pretty hopeless. Perhaps the best attempt to provide one would be to say

(i) that it is part of the concept of 'or' that '*A* or *B*' is true if and only if either '*A*' is true or '*B*' is true,

and

(ii) that it is part of the concept of 'not' that '*not A*' is true if and only if '*A*' is not true;

putting these together,

(iii) '*A* or *not A*' is true if and only if either '*A*' is true or '*A*' is not true. So either way, '*A* or *not A*' is true.

---

<sup>37</sup>Hilary Putnam's suggestion that quantum-mechanical results could be seen as indirect evidence against the distributive law is no counterexample, for two reasons. The main reason is that his claim that we could handle problems in quantum mechanics by giving up the distributive law was woefully under-supported. But in addition, he gave no support at all for his claim that the role of the quantum-mechanical facts would be to serve as evidence against the distributive law, rather than as an empirical trigger to motivate re-thinking whether there was a conceptual case for the distributive law. Perhaps the role of quantum mechanical evidence would be like the role of the discovery that there is no Santa Claus might have in the development of a free logic.

But of course the ‘so either way’ disguises the use of excluded middle, and the opponent of excluded middle should question its use here.

Another attempt to provide a “metaphysical justification” for the laws of classical logic would be to say that such justification somehow flows from the meanings of our logical words, where meanings are understood in terms of conceptual roles. What this amounts to is simply that any revision of certain logical laws is deemed a change in the meaning of some of the component logical concepts. But even if we grant the claim about change of meaning, it’s hard to see how it would provide justification in any interesting sense, given that the “old meanings” may be defective and *need* changing. For instance, if the laws of excluded middle and non-contradiction are built into the meaning of ‘not’, then it’s hard to see why the equivalence between ‘true( $\langle p \rangle$ )’ and ‘p’ isn’t built into the meaning of ‘true’. But keeping the principles built into *both these meanings together* is totally unacceptable because it leads to triviality via various semantic paradoxes.

Does the failure of these attempts at justification mean that the proponent of excluded middle needs a better justification? Well, it depends what you mean by justification. If you mean that the classical logician needs to respond to arguments that life would be better if we abandoned the law of excluded middle—e.g. that we could better account for vagueness, and have an account of truth that better handled the paradoxes—then yes, that kind of justification does seem required. This is a kind of dialectical justification, one needed *as a response to an opposing view*: it is a matter of arguing against the advice of the opponent of excluded middle. But what some who write on philosophy of logic seem to advocate is a different kind of justification, a metaphysical justification not directed against serious proponents of alternative logics. I think a defender of classical logic should reject the demand for that sort of justification. After all, it’s hard to see how to give a “metaphysical justification” for *any* logical law.

I don’t want to exaggerate the effect that rejecting the need of “metaphysical justification” would have on epistemology. Many distinctively philosophical issues in epistemology would still remain, or arguably so. Consider the Benacerraf (1973) worry about (certain kinds of) platonism in the philosophy of mathematics.<sup>38</sup> The worry is that the same

---

<sup>38</sup> If ‘platonism’ is defined simply as the belief that there are mathematical entities that aren’t

reasons that would lead us to advise against having beliefs about the happenings in a remote village in Nepal, when one has reason to think that there is no possible explanation of the reliability of those beliefs, should equally lead us to advise against having beliefs about mathematical entities platonistically construed, given that it appears that there is no possible explanation of the reliability of those beliefs. Maybe the apparent analogy here is no good—e.g., some have argued that because mathematics consists of “necessary truths”, explaining the reliability of the beliefs just amounts to explaining why we have those particular beliefs, which is easy. For present purposes there is no reason to take a stand on the quality of the argument; my present point is simply that whether the argument is any good seems independent of issues about “metaphysical justification”. In this it seems very different from issues about inductive skepticism.

### References

- Benacerraf, Paul 1973. “Mathematical truth” (*J. Phil.* 70: 661-79).
- Blackburn, Simon 1993. *Essays in Quasi-Realism* (Oxford University).
- 1998. *Ruling Passions: A Theory of Practical Reasoning* (Oxford University).
- Boghossian, Paul 2001. “How are objective epistemic reasons possible?” (*Phil. Stud.* 106: 1-40).
- 2006a. *Fear of Knowledge* (Oxford University).
- 2006b. “What is relativism?” In P. Greenough and M. Lynch, *Truth and Realism* (Oxford University).
- Chrisman, Matthew 2007. “From epistemic contextualism to epistemic expressivism” (*Phil. Studies* 135: 225-54).
- Dreier, Jamie 2004. “Meta-ethics and the problem of creeping minimalism” (*Phil. Persp.* 18: 23-44).
- Dworkin, Ronald 1996. “Objectivity and truth: you’d better believe it” (*Phil. and Publ. Aff.* 25: 87-139).
- Egan, Andy forthcoming. “*De gustibus non disputandum est* (at least not always)”, in R. Feldman and T. Warfield, eds., *Disagreement* (Oxford University).
- Field, Hartry 1982. “Realism and relativism”, (*J. Phil.* 79: 553-67).

---

physical, mental, etc., then I don’t think that the Benacerraf argument applies to all forms of platonism: see Field 1998a. But it may be unnatural to regard the positions it doesn’t apply to as forms of platonism.

- 1994. “Disquotational truth and factually defective discourse” (*Phil. Rev.* 103: 405-52; reprinted with significant corrections in Field 2001.)
- 1998a. “Mathematical objectivity and mathematical objects”, in S. Laurence and C. Macdonald, eds., *Contemporary Readings in the Foundations of Metaphysics* (Wiley-Blackwell).
- 1998b. “Epistemological nonfactualism and the a prioricity of logic” (*Phil. Stud.* 92: 1-24).
- 2000. “A priority as an evaluative notion”, in P. Boghossian and C. Peacocke, eds., *New Essays on the A Priori* (Oxford University).
- 2001. *Truth and the Absence of Fact* (Oxford University).
- 2005. “Recent debates about the a priori”, in T. Gendler and J. Hawthorne, eds., *Oxford Studies in Epistemology 1*.
- 2009. “What is the normative role of logic?” (*Proc. of Arist. Soc. Suppl.*).
- Gibbard, Allan 1990. *Wise Choices, Apt Feelings* (Harvard University).
- 2003. *Thinking How to Live* (Harvard University).
- 2008. “Rational credence and the value of truth”, in T. Gendler and J. Hawthorne, eds., *Oxford Studies in Epistemology 3*.
- Harman, Gilbert 1996. “Moral Relativism”, in G. Harman and J. Thomson, *Moral Relativism and Moral Objectivity* (Blackwell).
- Kölbel, Max 2004. “Indexical relativism vs. genuine relativism” (*International J. of Phil. Studies* 12: 297-313).
- Lewis, David 1971. “Immodest inductive methods” (*Phil. Sci.* 38: 54-63).
- 1974. “Spielman and Lewis on inductive immodesty” (*Phil. Sci.* 41: 84-5).
- 1986. *On the Plurality of Worlds* (Blackwell).
- MacFarlane, John 2005. “Making sense of relative truth” (*Proc. Arist. Soc.* 105: 321-39).
- forthcoming. “Non-indexical contextualism”. (*Synthese*).
- Pollock, John and Joseph Cruz, 1999. *Contemporary Theories of Knowledge* (Rowman and Little).
- Street, Sharon forthcoming. “Objectivity and truth: you’d better rethink it”.