

Epistemicism and Semantic Plasticity¹

John Hawthorne

(Very much work in progress)

\$0

I shall endeavor to make vivid a kind of puzzle that arises when Timothy Williamson's epistemicist machinery is applied to borderline cases of (i) personhood and (ii) semantic properties. My aim will be to make trouble for Williamson's development of the epistemicist view, and then to propose an alternative way of thinking about epistemicism.

\$1

Consider a Sorites series in which a subject *S* has his hair removed, one hair at a time, beginning with a full head of hair, ending with no hair at all. At the beginning, *S* is clearly not bald. At the end, *S* is clearly bald. However, there will be times when *S* is neither clearly bald, nor clearly not bald: at those times, *S* is a borderline case of baldness. Williamson's epistemicism combines the following theses about borderline cases of baldness:

- (1) The relevant instances of excluded middle hold. Supposing that our subject is now a borderline case of baldness, it is nevertheless true that

S is bald or ~ S is bald

- (2) Bivalence holds for any baldness ascription to *S*. Thus, whether or not *S* is now a borderline case of baldness

'S is bald' is true or 'S is bald' is false

- (3) If *S* is a borderline case of baldness, then we are unable to know whether or not *S* is bald.
- (4) Not all ignorance is due to vagueness. In the case of borderline cases, vagueness has a distinctive source, namely: if we had used the word 'bald' ever so slightly differently, we would have picked out a different property by 'bald'. We are insensitive to the ways that slight differences in usage make a difference to the semantic value of our terms. When ignorance is due to that kind of insensitivity, we have ignorance that is due to vagueness.

How does the kind of insensitivity in (4) make for ignorance? Well suppose 'bald' is true of *S*, but it is also true that if we had used the term 'bald' ever so slightly differently, 'bald' would have been false of *S*. If we are insensitive to the ways that slight differences in usage makes for a difference in semantic value, then insofar as we actually believe *S* is

¹ I am grateful for conversations with Cian Dorr, Kit Fine, Hud Hudson, David Manley, Ted Sider, Ryan Wasserman, Peter Van Inwagen, Dean Zimmerman and especially Timothy Williamson.

bald, there will be close worlds in which we make a mistake in a relevantly similar situations. Thus even if we believe that S is bald and get it right, we do not *know* that S is bald.

In a borderline case of baldness, then (a) there are a plurality of candidate semantic values, where a semantic value is a candidate for 'bald' insofar as, for all we are able to know, it is the actual semantic value of the term 'bald' (b) one of the candidates is the actual semantic value, which in turn determines the actual truth value of the relevant baldness ascription and (c) some of the candidates hold of the case at hand, others not.

(A bit of terminology: Let us say that a term is "semantically plastic" when (a) slight differences in usage make for differences in semantic value and (b) we are insensitive to the ways in which difference in usage makes for a difference in semantic value.)

\$2

Assuming that persons are material beings, the term 'person' appears to be vague.

(i) To begin: there is vagueness as to whether a person comes into existence. Even if, for example, one is convinced that a person begins with conception, there will be vagueness at the beginnings of personhood owing to vagueness as to when conception actually takes place (think of the trajectory of sperm, slowly approaching and entering the egg. It is clearly a vague matter when conception occurs). In general, all views about the beginnings of personhood use vague predicates in the favored criterion – even granting any given one of those views, vagueness along the temporal dimension will not disappear.

(ii) And unless one believes that persons enjoy life (or at least existence) everlasting, there will be vagueness as to when a person's existence comes to an end. Once one remembers that such predicates as 'dies' and 'is braindead' and so on are vague, the relevant thesis about persons should be obvious enough.

(iii) Further, there is vagueness as to where the spatial boundaries of a person lie. There are for example, certain atoms in the vicinity of my surface such that it is a vague matter whether or not they are parts of me.

(And there is also vagueness as to whether various less sophisticated beings counts as people. And so on.)

\$3

Let us pick on a case. Suppose that there is a speck of dirt – call it Tony—such that it is a vague matter whether or not the sentence 'The person sitting down has Tony as a part' (hereafter S) is true. Let us try to accommodate the vagueness of the case using epistemicist machinery. What we should say, it seems, is that there are a multitude of candidate semantic values for the term 'person' such that the truth value of S differs according to which of those candidates is adopted as the interpretation of 'person'. The point presumably extends to personal pronouns. It will thus presumably also be a vague matter whether or not the sentence 'He has Tony as a part' (pointing at the person in the chair) is true (hereafter S2). And this will be because there are a range of candidate semantic values for 'He' such that the sentence differs in truth value according to which semantic

value is adopted. Suppose S2 is true. There will be a meaning that could very easily have been given to 'He' such that S2 is false.

This semantic picture invites us to posit a plentitude of overlapping objects in the vicinity of the chair.² Only one of them falls within the extension of 'person'. Only one of them is the referent of the personal pronoun 'He'. Insofar as the object in question has a Tony as a part, then S1 and S2 are true. But owing to semantic plasticity, there are a variety of candidate semantic values (v1... vn) of 'person' that each associate a particular object with the definite description 'The person in the chair' (that is, relative to any candidate interpretation of 'person', a particular object will count as satisfying the definite description 'The person in the chair'). While some of the candidate semantic values associate an object containing Tony as a part with the definite description 'The person in the chair', others will not.

Let us focus on two of the candidates, one containing Tony as a part, the other not. Call them Grubby and Clean. Suppose Grubby and not Clean falls within the actual extension of the term 'person' in English (though of course we would be unable to know this). Then S1 and S2 are both true. But there is a possible tribe that uses the word 'person' ever so slightly differently such that Clean and not Grubby falls within the extension of the term 'person' in their mouth and S1 and S2 in their mouth are false. Such a tribe might even be actual. Pretend that there exists a tribe of Twinglish speakers that uses 'person' in such a way that 'person' is true of Clean and not Grubby. Then when a Twinglander says 'The person sitting in the chair has Tony as a part' he will express a false proposition even though we say something true.

There are a variety of overlapping object on the chair. Only one of them is a person. That is, only one of them falls under the extension of the actual semantic value of 'person'. Same, presumably, for such predicates as 'thinks' 'talks' and so on. Only one of the objects thinks. Only one of them talks. The one that is the person is also the one that thinks and talks. Other of the overlapping objects fall within the extension of candidate semantic values for 'thinks' 'talks' and 'is a person' (in the sense of candidacy explained.). Suppose Grubby thinks and talks. If our use of 'think' and 'talk' had been ever so slightly different in certain ways then 'think' and 'talk' would apply to Clean. Let

² There is an alternative conception that I will not pursue here. One might say in these cases that vagueness in spatio-temporal temporal boundaries does not have its source any vagueness associated with the term 'person' (and in the associated personal pronouns), but rather in vagueness associated with other pieces of vocabulary. Suppose it is a vague matter when I came into existence. It might be claimed that there is some particular object o, such that it is definite that 'person' is true of o, and definite that 'I' in my mouth refers to o, but that the location of o is indefinite owing to vagueness associated with the term 'occupies'. Now in the case of mereologically complex objects such as myself (assuming the falsity of Cartesianism), it seems clear that the location of the whole is derivative upon the location of the small parts. It would thus be strange to suppose it s a sharp matter what atom is or is not part of a mereologically complex thing at each time but that the location of the thing is vague, unless there is vagueness in the location of particular atoms. The latter is not plausibly the source of indeterminacy in my boundaries. So as far as I can tell, then, the current suggests will require positing vagueness in the *part of* relation: it is vague for various pairs xy that, as some given time, x is part of y. Assuming semantic plasticity, this will require that we posit a variety of candidate semantic values for 'is a part of', that differ intensionally. This strikes me as a rather radical tack, almost as radical as the idea that 'exists' is vague. I note further that no similar strategy is available for avoiding the analogous puzzles that arise in connection with semantic properties.

us say that an object thinks* iff it falls within the extension of one of the candidate semantic values for 'thinks'. We could similarly introduce the predicates 'person*' 'thinks*' and 'talks*'. There are many objects that on the chair that are persons*, which think* and which talk*. But only one of them thinks, talks and is a person.

Why do I insist that only one of the object in is a person? Well I take it that a feature of our usage we do not allow that many objects at a time are people when those objects mereologically overlap almost entirely. If semantic values are going to respect that aspect of usage then each candidate semantic value for 'person' will only allow one of the objects on the chair to fall within its extension.³

What makes, say, Grubby and not Clean the thinker? If we could know the answer to that question then (says the epistemicist) it would not be a vague matter whether Tony is part of the person. We cannot know what it is about our use of 'thinks' that determines one of the candidate semantic values to be the actual one. Our knowledge of semantic relations is incapable of extending that far. And that is precisely why ignorance arises in the case at hand.

Call the approach just sketched the 'Simple Epistemicist Treatment of Persons'

\$4

There is a problem for the simple epistemicist treatment. Let me illustrate it by an example. Suppose a Twinlander is sitting in a chair. Suppose that the Twinglander uses 'thinks' and 'person' ever so slightly differently (and is otherwise very much like an ordinary English speaker), so that the semantic value for 'thinks' and 'person', as used by the Twinglander, is different from ours. In particular, let us suppose that Grubby and Clean are both on that chair, and our use of 'person' is such that it is true of Grubby and not Clean and the Twinglanders use of 'person' is such that it is true of Clean and not Grubby.

Here are some very obvious truths

- (1) The Twinglander is the person sitting on the chair
- (2) The person sitting on the chair is only thing on the chair that is able to talk and think.
- (3) When the person sitting on the chair says 'I' the person is referring to himself.
- (4) If the person sitting on the chair says something of the form 'a is F' then that claim is true iff the predicate 'F' is the mouth of that person is true of the thing referred to by 'a'.

Let us add to these obvious truths the added facts provided by our epistemicist-driven description of the scenario

- (5) Grubby is the person sitting on the chair (and Clean is not)
- (6) 'is a person' in the mouth of the Twinglander is true of Clean and not Grubby.

³ Even one goes against that aspect of usage and counts many of the objects sitting on the chair as a person, some of the issues that follow will arise. For even if there are many persons on the chair, there will presumably be certain objects for which it is vague whether or not those objects are persons at all. If so, then apply the epistemicist machinery to those objects. The considerations that follow will now weigh heavily.

Suppose the Twinglander says 'I am a person'. We can deduce (i) that it is Grubby that says 'I am a person' (ii) that nothing else on the chair says 'I am a person'. (iii) that Grubby is referring to himself (iv) that 'I am a person' as uttered by Grubby is true iff 'is a person' in the mouth of Grubby is true of Grubby (v) 'is a person' in the mouth of Grubby is not true of Grubby. All of this has us conclude that when the Twinglander says 'I am a person', the Twinglander expresses a false proposition and that nothing says something true. The same argument, mutatis mutandis, could have been run for 'I think' and 'I talk'. We should now conclude further that if our use of 'person' and 'think' had been ever so slightly different, then the sentences 'I think' and 'I talk' would be false. Worse, if some of the candidate semantic values for 'think' are such that 'I think' comes out false, then we do not know whether or not 'I think' in our mouths is true. (At close worlds we make a mistake, which on a safety-driven conception of knowledge, undermines knowledge at the actual world.) Clearly, something has gone terribly wrong.

Let us get a bit clearer about the source of the problem. As things have been set up, it is our standards for 'person', 'thinks', 'talks' and so on that determine which of the candidate objects is an object that is self-referring, but it is an object's own, potentially different, standards that determines the extension of 'is a person', 'thinks' 'talks' and so on in its own mouth. It is our own standards that determine which objects are objects that are capable of engaging in the activity of drawing boundaries. In short, our standards determine which objects are boundary drawers. But it is the potentially different standards of the boundary drawer that determines where the extension of 'thinks' and 'person' in its mouth are to fall. Supposing the set of boundaries corresponding to 'person' that are drawn by a boundary drawer do not include its own boundary. Then were that boundary drawer ever to self-ascribe that predicate to itself it would make a mistake.

Let us say that a person uses a close variant of 'person' iff the semantic value of 'person' in the mouth of that person is a candidate semantic value for 'person' in English. Assuming that 'person' is semantically plastic, it seems very easy for a variant of 'person' in the mouth of a boundary drawer to be false of the boundary drawer. That is just to say it is easy for a person to be such that they use a close variant of 'person' in a way that is false of that person. Same for 'thinks'. The trouble is that we do not want to say that there are close variants of 'thinks' and 'person' such that those predicates are falsely self-ascribed. For it is all too easy to self-ascribe those predicates – 'I' thoughts will do the trick.

The simple epistemicist model needs supplementation or revision.

\$5

One response is to deny that self-ascription is easy for close-variants of English. One might insist that when the Twinglander says 'I think', the Twinglander is not referring to himself by 'I' but is referring to, say, Clean. On this model 'I think' in the mouth of the Twinglander is much more like 'He thinks' than it may first appear. In brief: whenever a

person uses a close variant of ‘thinks’ in such a way that the person does fall within the extension of ‘thinks’ then the person will have no readily available device for self-reference and in particular will not self refer when that person uses what would naturally be taken as a cognate of ‘I’.

This view may not be as bad as it first appears. It is here worth bearing in mind a potential symmetry between myself and the Twinglander. For on this view the Twinglander may well say ‘He cannot refer to himself’ where ‘He’ in the Twinglander’s mouth refers to something that mereologically overlaps me, but which is not identical to me, and that satisfies ‘person’ ‘talker’ ‘thinker’ and so on in *his* mouth. So the proponent of this view can do something to deflate the suggestion that we are really special by being able to self-refer. Nevertheless, I still think that the view will not strike us as appealing. There is something very strange about a view according to which at close worlds, many people (perhaps most people), do not have linguistic devices of self-reference. Relatedly, it is extremely natural to think that if a pronominal device has the conceptual role of the first person pronoun in a person’s cognitive life, then that pronoun will be a device of self-reference. The thought is a little rough and ready, owing to the rough and ready nature of the concept of “conceptual role”, but has some force nevertheless. “I” thoughts in the Twinglanders belief-box will have stereotypical roles in practical reason and so on that make it utterly natural to suppose that they are devices of self-reference.

§6

Return to the case of the case of the Twinglander. Suppose that by the standards of the Twinglander, ‘person’ is true of Clean. In short, Clean is the object that counts as the utterer of ‘I’- talk by the standards of the Twinglander. Perhaps the concept of a person is distinctive in that it always defers to the self-conception of people: an object can only count as a boundary drawer insofar as it draws its own boundaries at its own boundaries.⁴ On the hypothesis that Grubby is a person, Grubby counts Clean but not Grubby as the referent of ‘I’. That counts as a reductio of the idea that Grubby is a person at all.

Example: Suppose there is an Englander and a Twinglander in a chair. The intrinsic environment is pretty much the same for each. There are, inter alia, two objects Grubby_E and Clean_E in the Englander’s chair and two objects Grubby_T and Clean_T in the Twinglander’s chair. Suppose the Englander self-descriptions (without him knowing it) privileges Grubby_E and the Twinglander self-descriptions (without him knowing it) privileges Clean_T. Then ‘person’ in the Twinglander’s mouth is true of Grubby_E and Clean_T. Same for the Englander. Each defers to the other’s self-description as the prime semantic determinant of which object is the referent of ‘I’ thoughts and in term for which object counts as the extension of ‘person’. (I am here abstracting away from issues connected to what it may or may not take to be a person beyond being a thinker that can self-refer) .

⁴ This is not to say, of course, that a person cannot radically misdescribe himself. He may think himself an immaterial being when in fact he is material. And so on. The point is that we must allow questions of who the person is to march in step with questions about what is picked out by “I” in the person’s mouth. Thus once we concede that, overall, the pattern of usage in the person’s mouth privileges x over y as the reference of ‘I’, it is no longer an option to nevertheless reckon y as the person/thinker/utterer.

Our simple epistemicist position claimed, in effect, that people could very easily have used the term person in slightly different ways such that they did fall under the extension of ‘person’ in their mouth. Our revised position denies this.

\$7

Let us turn to the diachronic case. Suppose I begin life with the self-conception of a Twinglander and towards the end of my life move towards the self-conception of an Englander. (Of course, the shift may not be epistemically obvious to me: in fact if the shift is around borderline cases, it will not be.) Suppose my earlier self is sitting in a chair and my earlier self’s usage privileges Grubby (where now let Grubby be an object that always has Tony as a part) and not Clean (which never has Tony as a part) as the referent of ‘I’ thoughts. My later self however privileges Clean but not Grubby as the referent of ‘I’ thoughts. The natural application of the deferential conception is the following: My earlier self refers to Grubby with his ‘I’ thoughts and my later self refers to Clean with his ‘I’ thoughts. But this can’t be right. My earlier self IS my later self. But Grubby is not identical to Clean. When I look back on my earlier self, I want to say ‘I was referring to myself when I said ‘I am hungry’. If I am Clean then it cannot be that my earlier self refers to Grubby. The problem is analogous to the earlier one. The ascriber’s usage puts semantic pressure to count one thing as the referent of its ‘I’ thoughts, whereas the ascriber’s standards on who is to count as a thinker in the first place puts semantic pressure towards a different thing to count as the referent of ‘I’ thoughts. Where the target and the ascriber take themselves to be one and the same person, the ‘to each his own’ deferential strategy cannot be made to work. At this point the epistemicist needs to appeal to a new, anti-individualistic, theme.

Return to the case of ‘bald’. It is wrong to think that the extension of ‘bald’ in my mouth is simply a matter of how I use ‘bald’. My own use creates various semantic pressures, but I am a member of a linguistic community. The usage of others also contributes to the extension of the term in my mouth. Indeed, one reason – though perhaps not the only one – as to why I cannot know which value is the semantic value of ‘bald’ in my mouth is the fact that I am not privy to all the details of others’ usage.⁵ Now what goes for my relation to others in the linguistic community may go for my later or earlier self. The extension of ‘person’ in my mouth – and relatedly, the referent of ‘I’ may be constitutively determined by the usage of my later or earlier self. Suppose, to simplify, a community consisted of two individuals A and B. A’s usage of ‘bald’ may favor a cutoff at 17 hairs. Roughly speaking: if A was the only member of the community, then the extension of ‘bald’ in his mouth would include all and only people with 17 hairs or less. Suppose, meanwhile, B’s usage favors a cutoff at 15 hairs. This does not mean that B and A have different semantic values for ‘bald’. The fact that they translate each other homophonically puts semantic pressure against such a resolution. If God were to interpret them he would likely say that their terms have the same semantic value, adopting some appropriate weighting of the various semantic pressures at play. We cannot know how the weighting would proceed of course – our lack of knowledge of the details of the relevant laws of semantics is what gives rise to the phenomenon of vagueness (says the epistemicist). Similar remarks apply, *mutatis mutandis*, to the case at hand. If earlier

⁵ Williamson ref.

usage favors Grubby, the later Clean, but there is considerable semantic pressure for uniformity of semantic value across uses, then the matter is resolved by some (we know not what) weighting of semantic pressures to yield uniform semantic value. Perhaps the usage that favors Grubby loses out to the usage that favors Clean. In that case, both the early and later self use 'person' in a way that the extension of 'person' includes Clean and not Grubby. In that case no utterance of 'I am a person' gets to be false: on both occasions, it is Clean that makes the utterance. Grubby is not a person at any time. Both utterances come out true since Clean is always, Grubby never, included in the extension of 'person'.

§8

I have sketched a picture. Does it fit with the barebones of epistemicist I began with? It is of course perfectly consistent with excluded middle, with bivalence, and with the thesis that borderline cases are beset by ignorance. What is less clear on this picture is that there is semantic plasticity in borderline cases. More specifically, it is not clear that the intension associated with 'person' could easily have been different. To make this vivid, consider a world *W* containing nothing but a Twinglander sitting in a chair. If 'person' is deferential in the way described and the Twinglander's usage favors Grubby, then I should say that Grubby is the person in that world. Could the intension of 'person' have easily been such that it did not deliver Grubby as the extension taking *W* as its argument? (I am operating here with a standard conception of intensions as functions from worlds to extensions.) It would seem that a word that was not deferential in the way described would not be a close variant of 'person'. But that inclines me to think that there is no close variant of 'person' that does not deliver the set containing Grubby as value taking *W* as argument. We want to say that it is a vague matter whether the person in *W* has Tony as a part. But it is not at all clear that the ignorance can be explained as a matter of semantic plasticity.

§9

(An aside: Suppose one adopts the epistemicist picture sketched thus far. Given the plenitudinous ontology in the background and the willingness to defer to self-conceptions as the determinants of boundaries, it is at least natural to extend the picture to allow that some people have wildly different boundaries simply on account of wildly different self-conceptions. To make this vivid, consider a community of Eggers. An Egger believes that he or she came into existence as a human Egg. An Egger will say that fertilization made him or her gendered (and in general a lot more interesting). But on an Egger's self conception, his or her existence began with an egg. Embrace the plenitudinous ontology suggested by epistemicism and it is very natural to at least suppose that there is an object – don't ask yet whether it is a person – which comes into existence at the time that the egg does, which endures for 70 or 80 years and which has all the intrinsic requirements for thought later in life. Run with the deferential perspective and I think it very natural to say that these are the objects that the Eggers refer to by their 'I' thoughts. Some people, then, come into existence prior to fertilization. Not us perhaps. But the Eggers do. Is this a reductio of the deferential conception of persons? I myself don't think so. Perhaps some will.)

The issues I have raised are far from unique to epistemicism. Consider the main competing theory of vagueness,⁶ supervaluationism. The supervaluationist uses supertruth and superfalsity as her primary concepts of semantic evaluation. A sentence is supertrue iff it is true on all precisifications, superfalse iff false on all precisifications. A borderline case is one where the relevant sentence is true on some precisifications, false on others. Suppose ‘He has Tony as a part’ is borderline (where ‘Tony’ is precise). That will be because there are various precisifications of ‘He’, some of which contain Tony as a part, others not. Here too we have a plenitude of objects required by the semantics. Suppose further we embrace the analogue of semantic plasticity: small shifts in use generate small shifts in the range of acceptable precisifications of a term. Suppose now a Twinglander has a slightly different set of acceptable precisifications for ‘person’ than I do. Now the competing pressures described above will arise. On the one hand, I want to say that the Twinglander self-refers by ‘I’. This encourages me to treat all and only the acceptable precisifications of ‘The Twinglander’ in my mouth as acceptable precisifications of ‘I’ in the Twinglander’s mouth. But suppose some of the acceptable precisifications of ‘The Twinglander’ in my mouth are not acceptable precisifications of ‘person’ in the Twinglander’s mouth. Then there is a threat that ‘I am a person’, in the mouth of the Twinglander, will not come out supertrue. (Same for ‘I think’). Such a sentence will have to be reckoned borderline. This result is intuitively unacceptable. The same theme is in play, namely competing pressures on the semantics of ‘I’ in the mouth of another, generated on the one hand by my conception of a person and on the other by the self-conception of the other. It is thus relatively straightforward to recast the issues just discussed within the alternative semantic framework of supervaluationism. A shift from epistemicism to supervaluationism will thus not make the problems go away, so long as semantic plasticity is maintained.

(I also note that one currently popular approach to vagueness does particularly badly in connection with our puzzle cases. Delia Graf and Scott Soames have advocated a “shifting sands” approach to vagueness⁷, according to which the boundary associated with a vague term actually shifts during the process of moving up and down a Sorites series (never mind the details). Suppose I am confronted with a Sorites series in which I am asked of a series of particles whether each is part of Hartry Field, where each case is further from Hartry’s center of mass than the previous one, and where some of the cases are in the vicinity of Hartry’s boundary. The shifting sands view would apparently have us believe that the extension of ‘part of Hartry Field’ shifts during the series of judgments. Apparently, then, ‘Hartry Field’ in my mouth, may well denote a larger object by the time that I am done! We thus appear to have surprising capacities to make people grow.)

⁶ Or at least until recently the main competitor. The contemporary landscape is more complicated, in no small measure thanks to some people here today.

⁷ refs

I now turn to the case of semantic predicates. Williamson is happy to suppose that his picture extends to such predicates as ‘refers’ and ‘is true’. Thus, in a reply to a commentary by Stephen Schiffer, he writes:

... semantic ascent preserves vagueness. For example, since it is clear that something is bald if and only if it is in the extension of ‘bald’, ‘bald’ has the same borderline cases as ‘in the extension of “bald”’.

My general explanation of the ignorance that constitutes vagueness extends to semantic terms. Although someone may judge truly ‘Baldness is the property of having fewer than 3832 hairs on one’s scalp’, the judgement does not express knowledge, for whatever produced a judgment in those words could very easily have done so even if the overall use of ‘bald’ had been very slightly shifted (as it could very easily have been) in such a way that it referred to the property of having fewer than 3831 hairs on one’s scalp, in which case the judgement then made in those words would have been false. What produces the judgement does not produce true judgments reliably enough to produce knowledge. ... To extend this explanation of our non-semantic ignorance to an explanation of our semantic ignorance, note that in the envisaged counterfactual circumstances the sentence “‘Bald’ refers to baldness’ naturally still commands assent (clearly, ‘bald’ refers to baldness). In those circumstances, the false judgement in the words ‘Baldness is the property of having fewer than 3,832 hairs on one’s scalp’ goes with a false judgment in the words “‘Baldness’ refers to the property of having fewer than 3,832 hairs on one’s scalp’. Although someone may use the latter words to make a true judgment in the actual circumstances, the judgement does not express knowledge, for what produces it does not produce true judgements reliably enough to produce knowledge. Thus the account explains equally why we are not in a position to know that ‘baldness’ refers to the property of having fewer than 3,832 hairs on one’s scalp.⁸

I think that this passage betrays a confusion that is easy to overlook. Let us distinguish semantic plasticity – the phenomenon whereby the intension associated with a term could easily have been different – from extensional plasticity – the phenomenon whereby the extension of a term could easily have been different. Suppose I am moody. I fall under the extension of ‘happy at noon, April 15, 2004’. Being moody, I could very easily have fallen under the extension of ‘not happy’. Thus, there are close worlds where the extension of ‘happy at noon, April 15, 2003’ is different to what it actually is. But *that* phenomenon obviously does not by itself indicate that ‘happy’ is semantically plastic. Now if ‘bald’ is semantically plastic in the way that Williamson envisages, that certainly means that, say, ‘expresses’ (conceived of as a relation between a noise token and a property) is extensionally plastic. A noise token may in this world express a property P, baldness, and yet at a close world that noise token not express that property. Similarly, if we allow ‘c means p by noise type n’ to express a three place relation

⁸ ‘Reply to Commentators’ PPR December 1997, Vol LVII No. 4, p. 947-8

between a phonetic or graphemic type, a community and a property, then the semantic plasticity of 'bald' will make for extensional plasticity with regard to that ternary predicate. Moreover, it is plausible enough to think that such extensional plasticity, coupled with our insensitivity to the ways that slight shifts in use make for differences in semantic value, will undermine the possibility of knowledge of the propositions expressed by the relevant metalinguistic claims in borderline cases. The words "Baldness' refers to the property of having fewer than 3,832 hairs on one's scalp' could easily have expressed a falsehood. And given our insensitivity to the shifts in semantic value, one who accepted that sentence would not plausibly express knowledge thereby. But none of this shows that semantic terms are themselves semantically plastic (as well as extensionally plastic). And thus none of this shows that ignorance in borderline cases can be traced in every case to semantic plasticity besetting some term figuring in the indeterminate claim.

\$12

We have seen that the considerations adduced by Williamson in the quoted paragraph do not demonstrate that semantic terms are semantically plastic. But is there any positive reason to think that they are not? Interestingly, the puzzle adduced earlier can be reproduced here. We generated havoc earlier by allowing our standards for what counts as a person to draw boundaries in different places to cognate terms (terms used at close worlds with almost indistinguishable conceptual roles) used by various counterfactual people themselves. What happens if we allow, say, the predicate 'true' to draw boundaries in a way that fail to match the boundaries drawn by people at close worlds who use a term with a conceptual role that bears the hallmark of our use of 'true'? Havoc similarly results.

Thus let us suppose that 'true' is semantically plastic, so that its intension at close worlds differs from its actual intension. Here, I am thinking of 'true' as a predicate of utterance tokens, so that its semantic value will be a function from worlds to sets of utterance tokens.⁹ Suppose then, that at a nearby world W the semantic value of 'true' was slightly different, so that each utterance of a particular sentence S by a particular community C fell under the extension of the semantic value expressed by our term 'true' (given W as argument), but did not fall under the extension of the semantic value of 'true' as used at the nearby world. Let us assume, as required by the Williamson picture, that at that world the use of 'true' is only ever so slightly different, so that the fundamental features of its conceptual role at our world – in particular the behavior gestured at by the "T-schema" – are intact.

Consider now an utterance of

'S' is true iff S

⁹ The points that I am making can, however, be recast in terms of other different semantical frameworks, including, for example, one that takes 'true' as a predicate of utterance types -- where an utterance type is individuated by a combination of phonemic/phonetic considerations and a specified community of users.

made by an inhabitant of the nearby world under consideration. By hypothesis the right hand side of the biconditional is true. How about the left hand side? By hypothesis, 'S' does not fall within the extension of 'true' as used by the community at that world. Thus, the community would be saying something false by the left hand side. Thus, if the community were to utter the relevant biconditional, the left hand side would be false, the right hand side true. The biconditional would be false. Assuming semantic plasticity, we have been led to conclude that at close worlds, certain counterparts of the T-schema are false! This seems just as bad as conceding that at close worlds people do not self-refer by 'I'. Note in both cases we supposed that a certain conceptual role is accompanied by a certain semantic achievement: a pronoun with the conceptual role of 'I' has self reference, a predicate with the conceptual role of 'true' will yield true instances of the associated T-schema. In both cases, semantic plasticity induces a detachment between the relevant conceptual role and the associated semantic achievement.

As in the case of 'I', one might try to soften the blow. "After all," it may be said, "while counterparts of the T-schema are false at close worlds, they are true*, where the property of being true* is the property expressed by 'true' at close worlds." But I take it that this is not satisfactory. Truth is the norm by which we evaluate both our actual and counterfactual selves. The response requires us to think that at nearby worlds truth doesn't really matter. As such, it is not acceptable.

Untoward results can also be reproduced for 'refers' 'expresses' 'designates' and so on. Suppose, say, that 'refers' is semantically plastic, so that while tokens of some counterfactual name *n* refer to *x*, the pair $\langle n, x \rangle$ does not fall under the extension of 'refers' as used at that (nearby) world. Consider now the claim

'*n*' refers to *n*

as used at that counterfactual world. That claim is true just in case the pair picked out by the flanking singular terms falls under the extension of the binary predicate. The referent of "'*n*'" is the name itself. By hypothesis the referent of '*n*' is *x*. By hypothesis the pair $\langle n, x \rangle$ does not fall under the extension of 'refers', as used by members of the counterfactual community under consideration. Thus certain instances of the "disquotational schema for reference" come out false at nearby counterfactual worlds. Once again, an intolerable result.

The lesson, I take it, is that we should be very cautious about positing semantic plasticity for semantic vocabulary. Not only does the quoted passage by Williamson fail to provide any reason for embracing it; they are excellent reasons for rejecting it.

\$13

Insofar as we are sympathetic to epistemicism, we are left with a residual problem. What exactly is distinctive of the ignorance due to vagueness? I have argued that it is implausible that our ignorance concerning the boundaries of personhood can be traced to semantic plasticity. But this does not seem to be a good reason for denying that in some very reasonable sense, 'person' is vague. We should similarly allow that in some very reasonable sense, certain claims of the form

‘bald’ is true of people with less than N hairs

are borderline, even though the vagueness of such claims cannot be traced to the fact that certain terms occurring in them are semantically plastic. (Even if ‘bald’ is semantically plastic, that does not mean that ‘bald’ is.) Now we have noted, of course, that semantic terms may well be beset by extensional plasticity in borderline cases. But we cannot say that in general, ignorance due to extensional plasticity makes for the kind of ignorance associated with borderline cases. Suppose a particle moves rapidly between point A and B, so rapidly that we cannot in principle discern whether, at a given time, the particle is at A or B. Consider the claim ‘The particle is at A at noon’. There is extensional plasticity, sure enough. Suppose ‘is at A at noon’ is true of the particle. That predicate could easily have been false of the particle. But this case does not in any way have the feel of a case in which there is ignorance due to vagueness.

So let us reexamine the question as to what the epistemicist should say about the ignorance that is distinctive of borderline cases. Let us begin with what might be called a picture of the epistemicist metaphysics of semantics. It would be very strange indeed to deny that semantical facts (and propositional attitude facts) supervene on a groundfloor comprised of a certain distribution of fundamental properties across space-time (which will be microphysical, assuming that some broad naturalistic picture is correct). The epistemicist is thus happy to believe that there is some sort of function from fundamental distributions to semantical facts. Call that function F. Meanwhile, semantical ignorance about a certain noise type may have two different sources. On the one hand, we may be ignorant of various facts about the groundfloor which serve as input to F. Such facts will, let us suppose (or, if you prefer, pretend), straightforwardly encode this or that fact about how the noise type is used some member of the community, by fellow members of the community, the causal relations of that noise type to this or that feature of the world and so on. Let us call this source of ignorance about semantic facts *groundfloor ignorance*. On the other hand, we may have an rather incomplete grasp of F itself, so that even if one were (idealizing now) to have a full grip on the array of fundamental facts, one would still not be in a position to discern the semantical facts on the grounds that one’s grasp of how the latter depends on the former is radically incomplete. Let us imagine that the nature of F could be captured by a set of semantical laws that describe how semantical facts depend on the groundfloor. Insofar as we didn’t know what the semantical laws were, we would have ignorance not traceable to groundfloor ignorance. Let us call the second kind of ignorance *semantico-nomic ignorance*.

Now it is quite clear that in a borderline case, Williamson supposes the ignorance not merely to be rooted in groundfloor ignorance: even if one knew all of the relevant groundfloor facts, one would not be able to make the ignorance go away. Supposing that the groundfloor facts are captured by P and the relevant semantic facts about the extension of some predicate Q. The problem is not merely that we do not know that P. It is that we are in no position to know that $P \supset Q$, even though that material conditional is presumably a necessary truth. Groundfloor omniscience would not remove our insensitivity to the true semantic mechanisms.

Notice now that this picture provides a plausible epistemicist account of ignorance due to vagueness that does not proceed by way of semantic plasticity: in cases which we call “ignorance due to vagueness”, we have a sentence that expresses a

proposition P such that our principled inability to know whether P is rooted in semantico-nomic ignorance. Even if, say, some claim of the form

(1) Tokens of ‘big number’ as used by community C are true of any number great than 154

are not semantically plastic, we may have a principled ignorance of their truth value that is rooted in semantico-nomic ignorance. Hence our ignorance of (1) will count as “ignorance due to vagueness”.

One might worry that the picture just sketched disrupts a safety-based conception of knowledge according to which belief is knowledge just in case there is no danger or error – that is, no error at “close worlds”.¹⁰ Suppose someone dogmatically believed some claim S of the form (1) above. Clearly such a person would not know that S even if it were true. Williamson provides us with a vision of how semantic plasticity explains ignorance in borderline cases: Suppose someone were to dogmatically accept a borderline claim S. Even if S is true, then, owing to semantic plasticity, S would express a falsehood at “close worlds”. Thus, at close worlds, the dogmatist would make a mistake. His actual belief thus turns out not to be safe and so he The dogmatist does not know S. Eschew plasticity for S and no similar explanation is available. Someone who dogmatically accepted S would, it seems express a truth at close worlds and so, by a “safety” theoretic test, counts as knowing that S is true.

Is this a real problem for the current brand of epistemicism? I don’t think so. To use the preceding line of thought against that account is to presuppose an all too crude safety-theoretic account of knowledge. We all know that if someone dogmatically cleaves to Goldbach’s conjecture, that will not in itself secure knowledge. But, given that there seems to be no semantic plasticity in the relevant mathematical language, such a dogmatist would not be in error at close worlds. In that case, we are hardly inclined to use a crude safety based account as grounds for admitting that the dogmatist knows Goldbach’s conjecture after all (assuming that it is true). We instead refine our conception of what knowledge comes to. Similar remarks apply to semantico-nomic ignorance, *mutatis mutandis*.

\$14

Suppose I introduce a predicate ‘is a dommal’ by a pair of stipulations: Let ‘is a dommal’ be true of dogs and false of non-animals.¹¹ My stipulations do not settle whether
(2) Cats are dommals

Williamson accepts bivalence here. On his preferred picture, the truth value is determined by a default principle, the two candidates being, roughly,

P1: A sentence is false unless one has done enough to secure its truth

And

P2: A sentence is true unless one has done enough to secure its falsehood.

¹⁰ I am grateful for conversations with Williamson here.

¹¹ ref

Williamson seems to think that he knows which default principle is the true one¹², but such an epistemic stance does not seem very plausible to me. Do we really have access, a priori or otherwise, to the relative merits of P1 and P2? I reckon it better to combine bivalence with an admission of principled ignorance: we do not know (2) owing to semantico-nomic ignorance. But notice that here, too, semantic plasticity seems utterly irrelevant. While I do not claim to know P1, it does seem implausible that P1 would hold at the actual world but P2 at close worlds. Semantic laws have more modal stability than that. Thus it seems that if (2) expresses a truth, it does so at close worlds as well. Our semantic-nomic ignorance is not allied to semantic plasticity.

Consideration of P1 and P2 does bring another concern to light.¹³ On the epistemicist proposal currently being considered, the phenomenon of “of ignorance due to vagueness” just is the phenomenon of ignorance rooted in semantico-nomic ignorance. But our ignorance of, say, P1 itself is rooted in semantico-nomic ignorance. Do we then want to say that P1 is vague? It might seem that from an epistemicist point of view, there is no deep difference between P1 and, say the continuum hypothesis in mathematics. Isn't it misleading, then, to say that our ignorance of the former, but not the latter, has to do with vagueness?

In reply, I do not find it particularly surprising that the epistemicist will need to disrupt some intuitions. It would be dishonest to represent epistemicist as leaving everything where it is, or even as saying that our ordinary understanding of matters, while incomplete, is perfectly in order. After all, if we recognized the truth of epistemicism, we would not find Sorites arguments paradoxical in the least: rather, the major premise would strike us as obviously false. There *is* an epistemological phenomenon in the vicinity of what we call “vagueness”, namely the category of principled ignorance rooted in semantico-nomic ignorance. But that this should turn out to be the root of vagueness is something of a surprise, and that this in turn should induce some conceptual reorientation is to be expected. In particular, it is to be expected that the epistemicist should reckon our ignorance of (2) and P1 more akin to ignorance of, say, the continuum hypothesis than we had previously recognized. That epistemicism effects such conceptual disruption may, perhaps, be reckoned a cost of the view. But since it is arguable that any theory of vagueness induces some disruption or other, this charge against epistemicism is hardly decisive.

(Another illustration of the surprises in store: Suppose I define a ‘tulower’ by the pair of stipulations: if something is a tulip it is a tulower, and if something is not a flower it is not a tulower. Consider now (3) ~ Daisies are tulowers. Similar considerations as above incline us to the view that (3) is vague. However, it seems very clear that if P1 is in force, it applies to both (2) and (3) and that if P2 is in force it applies to both (2) and (3). It thus seems that we can know that the conjunction of (2) and (3) is false, and hence it turns out that the conjunction of (2) and (3) is determinately false..)

¹² ref

¹³ Thanks to Kit Fine here.

Let us turn to the phenomenon of semantico-nomic ignorance itself. The picture is one according to which semantic mechanisms transcend our grasp of them in a deep and principled way. Some will find this deeply intolerable. It is interesting, here, to note a contrast between our attitudes towards mathematics and semantics. In the realm of mathematics, the view that there are evidence-transcendent features of this or that mathematical structure, while hotly debated, is not regarded as extreme or bizarre. Yet analogous views about semantics are apt to strike readers as somewhat outrageous. This reaction is at least in part rooted in a reluctance to recognize semantical properties as natural kinds, joints in nature with distinctive real essences. This “hypeinflationary” conception of semantical properties would not, of course, suffice to establish the current brand of epistemicism. But it would render the idea of semantico-nomic ignorance rather more palatable and thus help to make my favored version of epistemicism a going concern.

Let me thus offer a few preliminary motivating remarks in support of hyperinflationism.

Consider first the following frequently voiced concern about epistemicism:

For any predicate, there are every so many functions from use to extension that “fit” the use of that predicate. What on earth could it be that makes one of those functions special in such a way that ‘true of’ should be specially associated with it? Shouldn’t we instead make every attempt to do justice to the thought that each of the functions provides an equally good candidate extension?¹⁴

The concern needs refinement. Recalling Kripkenstein,¹⁵ none of us (or hardly any of us) think that quus is an equally good candidate semantic value for ‘plus’ in the mouth of our earlier selves as plus (where plus and quus are functions that differ only with regard to pairs of natural numbers whose sum we are unable to entertain due to considerations of our finitude.) But both candidates “fit” use in some fairly obvious sense, since each interpretation is equally charitable with regard to our actual and counterfactual use of ‘plus’ (so long as suitable compensating adjustments are made in the interpretation of other pieces of arithmetical vocabulary in which generations about additions are stated.) The lesson generalizes to non-arithmetical vocabulary. Bizarre interpretations can be concocted to “fit” use which none of us are very inclined to think are acceptable interpretations.

In response to all this, some will go the way of Kripkenstein’s skeptical solution,¹⁶ combining a suitably disquotationalist story about our truth predicate with a recognition that there are no deep objective constraints on the acceptability of a translation. The Quine-Field development of this view would have us believe that an

¹⁴ The relevant notion of “fit” deployed in such arguments is typically left unexplained, but I take it that talk of an interpretation “fitting” use of some term is tantamount to a claim that some interpretation provides some reasonably charitable interpretation of our settled dispositions to use a term.

¹⁵ ref

¹⁶ ref

ascription of truth to some utterance made by my earlier self (or some interlocutor) has to be relativized to a translation scheme. Semantico-nomic ignorance will have no place in that framework. Indeed, the chasm between such theorists and the current brand of epistemicism is far too vast for me to hold out much hope of closing it here. (For many of us, it is cost enough for that view that it relinquishes all hope of salvaging straightforward truth for claims made by our earlier selves and our fellows.)

Of more interest to me here is the Lewisian reaction to Kripkenstein,¹⁷ one which allows the distinction between natural and gerrymandered properties to do work in the foundations of semantics. Roughly speaking, the picture maintains there are two desiderata on interpretation, namely: (a) The Requirement of Charity: *ceteris paribus*, interpret us so that our claims come out true so interpreted; (b) The Requirement of Eligibility: *ceteris paribus*, interpret us so that our predicates get assigned more rather than less natural properties as their semantic values.

At first blush, the Lewisian approach seems to close the door to Kripkenstein without offering much encouragement for epistemicism. What makes a plus-interpretation more acceptable than a quus interpretation? Well, while both interpretations may do equally well on the score of charity, one interpretation scores far higher with regard to eligibility. Thus the quus interpretation can be discounted. Consider by contrast the case of ‘bald’. Each “candidate” semantic value is, intuitively equally natural. So neither charity nor eligibility can break the tie. So it seems as if the epistemicist has little to gain from a Lewisian distinction between natural and non-natural properties.¹⁸

Such reaction would be far too hasty, I think. Let me explore a few themes in that connection, making vivid those ways in which Lewis can serve as a springboard for hyperinflationism.

(i) Suppose one were to embrace the Requirement of Eligibility, along with an objective distinction between natural and unnatural properties. This is to already to recognize the existence of deep principles about semantics that transcend the ken of ordinary folk. Perhaps one might think that we semantic theorists can appreciate the plausibility of such a requirement and even know that it is probably true. But it would be outrageous to suggest that ordinary linguistic competence brings with it knowledge of any such principle. To claim that such a principle is known “implicitly” by ordinary folk is to court further confusion: to claim that those principles that describe how terms refer are automatically known implicitly by people simply on account of their ability to refer is, in truth, no more plausibly than the claim that those principles that describe how we maintain our balance are automatically known implicitly by people simply on account of their ability to maintain their balance. The eligibility requirement does not govern the semantics of ordinary folk not by being known implicitly by them. Rather, it governs the semantics of ordinary folk (if it does so at all) by virtue of being a correct (if partial) account of the nature of semantical relations. Accepting the eligibility requirement is, obviously, not yet to accede to epistemicism. But to accept it is to embrace the existence

¹⁷ ref

¹⁸ Relatedly, Lewis would certainly have it that there is no uniquely best way to weight charity versus eligibility when it comes to assigning semantic value. If one interpretation scores slightly better on charity, another slightly better on eligibility, then both will likely stand as acceptable interpretations that can be supervaluated over.

of fundamental semantic mechanisms that are beyond the ken of ordinary folk, a move that should provide real encouragement indeed to the epistemicist.

(ii) The version of epistemicism I am interested in is best served by a metaphysics according to which semantic properties – reference, truth and so on – are themselves natural kinds, joints in nature. Where a property marks a natural kind, we are open to the thought that it has a “real essence” that transcends our ordinary understanding of it, even one that in some respects transcends our cognitive capacities. The fundamental metaphysical task of my epistemicist, then, is to render plausible the picture of semantic properties as joints in nature. Does the Lewisian metaphysics help or hinder in this way?

First, some preliminaries. Lewis embraces a plenitude of properties, some of which are metaphysically “haloed”, that is natural. More precisely, there is a continuum from more to less natural properties, with perfectly natural properties at one end, and increasing “gruesomeness” as one moves along the continuum. Which *are* the natural properties? Even supposing that we think that everything supervenes on physics, the issue is not settled. For if we accept a natural property framework, we must choose between an austere physicalism on the one hand and what might be called an “emergentist” framework on the other. According to the austere physicalist, the perfectly natural properties will only be found at the microphysical groundfloor, relative naturalness being a matter of definitional distance from the perfectly natural properties: to calibrate the naturalness of a property, see how complicated the definition of that property would be in a “canonical” language in which each predicate corresponded to a perfectly natural property.¹⁹ From such a perspective, the property of, say, being a chair will likely turn out hopelessly unnatural, far less natural than, say, the disjunctive property of being either a hydrogen atom or being fifteen feet from a quark. Indeed, it wouldn’t be surprising if, lacking functional predicates, the canonical definition of a chair was infinitary. The “emergentist” by contrast, believes that naturalness is not a matter of mere definitional distance from the microphysical groundfloor. Perhaps being a cat is far more natural than certain properties far more easily definable in Lewis’ canonical language. On the emergentist conception of things, there is no algorithm available for calibrating naturalness in terms of a perfect microphysical language.

Now Lewis’ own development of the eligibility view certainly provides a hindrance to the picture of semantical joints – joints delineated by semantic predicates themselves, since his physicalism is an austere one. Though lacking the space to develop the point here, I suspect that we find hereabouts a fundamental tension in his world view: on the one hand he wishes the eligibility requirement to do dispel the specter of rampant indeterminacy presented by Kripkenstein and Quine. Yet on the other hand, he offers us an austere physicalist account of what eligibility comes to. It does not seem that the two perspectives can be reconciled. How can the eligibility requirement provide some reasonable measure of determinacy for ‘gavagai’ if the property of being a rabbit turns out to be hopeless gruesome? Far better, it seems to me, to opt for an emergentist physicalism, in which semantical joints remain a live option.

We can go further. The eligibility framework offers a natural perspective on the presence and absence of semantic plasticity. Suppose there is a highly natural property that distinguishes itself among the properties that reasonably well “fit” the use of a predicate. The other “candidates” are far less natural and so the highly natural property

¹⁹ ref

easily wins the semantic competition. Even if the use of the term had been slightly different, the highly natural property would win the competition, since even a slightly lower score vis a vis a gruesome property on the score of charity would be trumped by a far higher score in naturalness. Using language that has recently become popular: the highly natural property serves as a reference magnet. But we have seen above that semantical predicates are not semantically plastic. The reasonable conclusion seems to be that semantical properties are reference magnets and therefore highly natural themselves. Epistemicism is not yet forced upon us, but a suitable metaphysical underpinning for such a view – one replete with semantic magnets -- is now in place.