4. The Epistemic Theory of Vagueness

So far we have looked at theories on which vagueness is a *semantic* phenomenon. We will now look at some views that locate the distinctive features of vagueness elsewhere, though don't think that we've stopped talking about semantic theories! Today we look at the epistemic view. On some interpretations, this view was defended by Stoic logicians, but I'll leave the correctness of that claim to people with better knowledge of history than I. In the late 1980s and early 1990s, Roy Sorensen and Paul Horwich defended it, and Sorensen has recently released a book outlining his version of the theory. But the theory's current prominence is due largely to its defence by Timothy Williamson, and we will concentrate on his work.

4.1. Argument from T-schema

So here's a quick argument that no sentence that says something is neither true nor false. The argument is in §7.2 of Williamson (1994). We assume that it is false, that there is a sentence *u* and a proposition *p* such that *u* says that *p*, and *u* is neither true nor false, and derive a contradiction.

*1.	*u* says that *p*	
*2.	It is not the case that: *u* is true or *u* is false.	
*3.	If *u* says that *p*, then *u* is true iff *p*	
*4.	If *u* says that *p*, then *u* is false iff $\neg p$	
5.	*u* is true iff *p*	MP 1,3
6.	*u* is false iff $\neg p$	MP 1,3
7.	It is not the case that: *p* or *u* is false.	Substitution 2,5
8.	It is not the case that: *p* or $\neg p$	Substitution 6,7
9.	$\neg p$ and $\neg\neg p$	DeMorgan 8

So we assumed that there is a sentence that says something that is neither true nor false, and two fairly plausible principles of truth. The only logical rules we appeal to in the proof are (a) modus ponens, (b) substitution - that from *A iff B* and …*A*… we can infer …*B*…, (c) an instance of DeMorgan's law, that from $\neg(A \vee B)$ we can infer $\neg A \wedge \neg B$ and, implicitly, a version of *reduction ad absurdum*, that if some assumptions imply a contradiction, then those assumptions are not all true. Most of these logical principles look fairly innocuous, the real issue is whether we are entitled to have the assumptions about truth. We have already commented on these principles a little above, so I will just make two comments here.

First, the principles used *here* are not *obviously* refuted by the existence of liar sentences. On *some* theories, the liar, *This sentence is false*, and even the truth-teller *This sentence is true*, do not express any proposition. Whether this position can be sustained, at least it isn't clear that 3 and 4 alone are inconsistent. This is an important point. This argument shows *at most* that 1, 2, 3 and 4 are inconsistent. If 3 and 4 are inconsistent, this doesn't tell us much about 1 and 2. But it is at least arguable that 3 and 4 are consistent, and hence at least arguable that they are true.

Secondly, the main reason for being doubtful that 3 and 4 are true is that they imply that 1 and 2 cannot both be true, but plausibly 1 and 2 can both be true. So the best argument *for* premises 3 and 4 here is to show that the intuitive argument for premises 1 and 2 fails. This is what Williamson spends the most effort doing, and it is to this question we now turn.

4.2. Gaps

Why might we think that a particular sentence containing vague terms is neither true nor false? Let us consider a particular sentence, say *Louis is bald*, where Louis is a penumbral case of vagueness. One (bad) argument for this conclusion goes as follows. It is obvious that *Louis is bald* is neither definitely true nor definitely false. Assume that a sentence is true only if it is definitely true, and false only if it is definitely false. This position is endorsed by supervaluationists, at least until higher-order vagueness gets considered. Then it follows that the sentence is neither true nor false. Of course, the assumption here stands in need of *some* justification.

Williamson points out that this argument might gain some plausibility if we equivocate over how we interpret *definitely*. We can either use this as a technical term, or as a natural language term. If we use it as a technical term, then we are entitled to stipulate that it connects with truth in the way specified. That is, we are entitled to stipulate that we intend *definitely* to express a concept of semantic definiteness, so that a sentence is true only if it is definitely true in this sense. But once we make that stipulation, we are no longer entitled to appeal to the 'obvious' fact that the sentence is neither definitely true nor definitely false. For our intuition that this is so is surely only evidence that it is so *if the terms are given their natural meaning*. And once we do that, it is not obvious that *definitely* should receive a semantic interpretation. Indeed, it is possible that it should receive an epistemic interpretation.

Well, this is possible provided we are a little cagey about what the particular epistemic interpretation is. Williamson says that we should read *definitely* as meaning, roughly, knowably. But nothing is knowable or unknowable *simpliciter* - things are only knowable or unknowable for a particular agent or class of agents. So if *definitely p* means *X can know that p*, then we have to ask, who is *X*?

It can't be the class of all agents. Presumably if *Louis is bald* is either true or false, then God could know which it is. Epistemicism isn't mean to imply atheism. (Indeed, Hud Hudson has been using epistemicism in his defence of an idiosyncratic, but seemingly *consistent*, version of theism.)

At the other end, it certainly can't just be me, or speakers around here now. I have no way of knowing what Caesar was thinking about when he crossed the Rubicon. Maybe that he should get around to paying Brutus back that money he lost betting on chariot races or there'd be hell to pay sometime soon. Maybe not. In any case, there are definite facts abou t what Caesar was thinking, although none of us are in a position to know what they are. (This is not to say that all the claims of the form *Caesar was then thinking about such-and-such* are definitely true or definitely false. There will be some indefinite cases, but they are a much smaller class than those about which we can have some knowledge.) Michael Dummett has occasionally flirted with the idea that we might be anti-realists about those aspects of the past about which we can have no knowledge, but I assume *we* would have no truck with this idea.

It can't be the class of all humans. There might be definite facts about what kind of qualia a rabbit would experience when being sucked into a black hole, but I doubt that any human could ever figure out what they are. I intend *definite* here to be used in its ordinary sense, so my evidence that there are definite facts of this sort is pretty much just my intuition that there are definite facts of this sort. Maybe the science of consciousness will advance far enough that we can know this, but I doubt it. For a more dramatic example, I doubt we could ever know whether there is an intelligent species whose entire career takes place outside our light cone. Maybe we could have fairly solid inductive evidence one way or the other, but it is possible that we would never *know*. Indeed, it's rather plausible that we would never even have decent evidence one way or the other. But there are still definite facts about whether there are civilisations outside our light cone, in a way that there doesn't seem to be a definite fact about whether Louis is bald.

Williamson aims to get out of these problems by analysing definiteness not as knowability for a certain class of entities, but as a certain kind of knowability. If you do not, and in fact could not given the limitations on your epistemic capacities, know that *s* does not mean p^*, and p^* is false, then *s* is not definitely true. You could not know that *s* is true for a distinctively *semantic* reason, you do not know that it does not mean some proposition that is false. This will involve more negations than we might be comfortable with, but we can give a formal definition of definiteness.

> *Definitely s* is true iff for all propositions *p*, if *X* cannot know that it is not the case that *s* says that *p*, then *p*

There is still some agent relativity because we do not specify who can go in place of *X*, but we ignore that from now on. If you are worried, let *X* range over actual humans. If *a* is a borderline case of being an *F*, then there will be properties O_1 and O_2 such that *a* is O_1 but not O_2, and *X* does not know whether *F* means O_1 or O_2. (We will have reason below to alter this definition, but it seems to be the definition Williamson endorses, though he never explicitly says so, and it will certainly do as a first approximation.)

4.3. Margin of Error Principles

So why can't we know just what property is named by 'bald'? Well, that's putting the question in a misleading form. We can know which property is named by 'bald', namely, baldness. What we cannot know, according to the epistemicist, is that 'bald' does not mean baldness*, where baldness* is a property very much, but not exactly, like baldness. This requires an odd construal of know-which claims, one at odds with the natural suggestion in "Whether Reports", but perhaps this isn't a major problem. Still, now that we've formulated the question aright, why can't we know that 'bald' does not mean one of these other properties?

Williamson suggests it is because knowledge is governed by certain 'm
In very rough form, the idea is that if *X* knows that *p*, then *p* must be true in all 'nearby' situations. This kind of principle obviously has a lot of intuitive support. It provides very natural answers when we consider, as Williamson often does, cases where *p* is about the height of a particular object. In these cases, we have a natural measure of nearness, but the principle retains its plausibility in cases where similarity is more nebulous.

This principle is what drives standard sceptical intuitions: the sceptic tries to convince us that situations where all our sensations are provided artificially are nearby in the salient sense. It is what drives the original Gettier case: Smith does not know a disjunction because in an obviously nearby case the disjunct for which he has no evidence is false. (The analogy between the sceptical reasoning and the Gettier reasoning should raise some flags, but this ain't an epistemology book, so we'll pass on that for now. In any case, I can never convince *anyone* about *anything* concerning Gettier cases.) The principle also does a lot of work in many of the post-Gettier cases, though we won't go through all of these here.

Assume, for the sake of the argument, that an adult American male satisfies *tall* in the language we are speaking now iff he is more than 179cm tall. This fact is determined by the pattern of usage of 'tall' in the linguistic community we inhabit. (It is important for Williamson's version of epistemicism that meanings are determined by communal languages rather than personal idiolects. We could have some fun playing around with this assumption, but we'll just accept it for now.) Since *tall* is not a natural kind term, it does not 'lock on' to this extension. Had the usage of *tall* been just a little different, then its extension would have been different. Perhaps, has we been just a little less generous in applying the term, then any adult American male over 179.5cm would have satisfied *tall*. This is a nearby case, by any reasonable measure. So we cannot know that the cut-off, the borderline between the tall and the not tall, falls at exactly 179cm. Even if we did *believe* this, which evidently we do not, our belief would not count as knowledge, any more than Gettier beliefs count as knowledge. Consider now a particular person who is 179.2cm tall. We cannot know whether *He is tall* says that he is over 179cm tall or that he is over 179.5cm tall, or some other proposition. So there is some proposition *p*, namely that he is over 179.5cm tall, and we do not know whether the sentence expresses that proposition, and that proposition is false. So this sentence is not definitely true, just as we might have hoped and expected.

4.4. Higher-Order Vagueness

The epistemicists' characterisation of vagueness appears to permit a very elegant treatment of higher-order vageness. Indeed, higher-order vagueness is really the point where epistemicists make up ground on their opponents. We already saw the problems that degree theories and supervaluational theories have with higher-order vagueness, so if epistemicism offers a step forward, this might count as a major advantage.

If we can have an *a* and an *F* such that it is indefinite whether *a* is definitely an *F*, then we have higher-order vagueness. And the epistemic theory of vagueness, plus the margin-of-error model of knowledge, promises to make all this possible. Imagine again that an adult American male satisfies the predicate *tall for an adult American male* (or, for short, *tall* where we assume context does the rest) iff he is over 179cm tall. Now because of our ignorance about the meaning of the words, we may know that the boundary is between 177cm and 181cm, but not know where it is within that area. So anyone below 177cm is definitely not tall, anyone above 181cm tall is definitely tall, and anyone between those two heights is a penumbral case.

Now just as we cannot know just where the tall/not-tall boundary is (i.e. at 179cm) we cannot know where the definitely-tall/not-definitely-tall boundary lies (i.e. at 181cm). Maybe we know that this boundary is above 179.5cm, and below 182.5cm, but we do not know precisely where it is. (We will come back later to the question of whether the margins of error in this case should be as large as in the original case, or, as I have assumed here, smaller than in the original case.) So someone who is exactly 182cm tall will be definitely tall, but we cannot know they are definitely tall, so they are not definitely definitely tall. Hence we have a borderline case of being definitely tall. Hence we have higher order vagueness, as required.

There are two natural objections to this picture, neither of which seems to be ultimately successful. The first is that it denies us a natural kind of privileged access to our knowledge states. The second is that the kind of ignorance appealed to in the generation of higher order vagueness is not *semantic* ignorance of the right kind, so the generation of a second-order borderline case does not go through. We'll look at these in order.

In the second-order borderline case, we have to have the following odd combination of facts. When faced with someone whom I know to be 182cm tall, I have to know that this person is tall, but not know that I know that this person is tall. This seems strange. One way to bring out the strangeness of it is to consider what I would do if asked whether that person was tall. So imagine the following dialogue.

Q: Is he tall? (Pointing at someone we know to be 182cm tall)
A: Yes.
Q: You know that do you?
A: Er, I'm not sure.

My first answer is correct, he is tall. Of course, there are more to the norms of conversation than just telling the truth. Someone who just makes lucky guesses is, in some sense, not a great conversational partner. (Though if *you* know they are *lucky* guesses, you might not think this person is too bad!) Ideally, we want people to tell us things they *know* to be true. But my first answer satisfies that constraint - I do know that he is tall. On the other hand, if I answer *yes* to the second question, I'll be saying something that I don't know to be true, that I know the person to be tall. There might even be something wrong with my second answer anyway, because it is possible that I don't know that I'm not sure that I know the person is tall. More generally, it seems that if *yes* was the appropriate answer to the first question, then *yes* should be the appropriate answer to the second question. And if this is true, and it is only appropriate to assert sentences we know to be true, then knowing something implies knowing that you know it.

The principle that you should only assert what you know does a fair bit of work here, and you might think it is inappropriate to appeal to it in a contentious argument. But the most prominent defender of that principle in contemporary philosophy is none other than *Timothy Williamson*. And this is not entirely coincidental. To the extent that epistemicists can say *anything* about the Sorites, what they say rests heavily on this principle. So we are entitled to appeal to it here. Williamson has to bite the bullet and say that it can be appropriate to answer *Yes* to the first question even if it is inappropriate to answer *Yes* to the second. That is exactly what he does, with an interesting Sorites-like argument to back it up.

What is at issue here is what has become known as the KK principle. Formally, this principle says that in epistemic modal logics, $\Box A \rightarrow \Box\Box A$ is an axiom. Just what this amounts to in non-formal languages depends on just how we interpret the box. It is common to hold that this principle is *false* if the box means 'It is known (by *X*) that', because *X* might know something without having reflected on the fact that she knows it, and hence without knowing that she knows it. However, this alone provides no reason to deny the KK principle when the box is interpreted as 'It is possible (for *X*) to know that'. Yet if we are to interpret 'definitely' as, roughly, 'knowably', and to deny that *Definitely s* entails *Definitely definitely s*, then we have to deny the KK principle on just this interpretation. Williamson has an argument that we should deny it in just these cases.

The argument turns crucially on the margin of error principle stated above. We said that if *X* knows that *p*, then it all nearby cases, it must be the case that *p*. Apply this to the special case where *p* is a claim about what *X* knows. So if it is possible for *X* to know that *X* knows that *p*, in all nearby cases it must be possible for *X* to know that *p*. Now assume that whenever it is possible for *X* to know that *p*, it is possible for *X* to know that *X* knows that *p*, i.e. that the KK principle holds. This implies that in all nearby cases, it is possible for *X* to know that *X* knows that *p*. Iterate the above reasoning, and we get that in all cases that are nearby to nearby cases, it is possible for *X* to know that *X* knows that *p*. Another iteration gives us that in all cases nearby to cases nearby to nearby cases, it is possible for *X* to know that *X* knows that *p*, and so on.

Now the problem is that there are obviously cases where *X* does know that *p*, and even know that she knows that *p*, even though there are cases nearby to cases nearby by to … nearby cases where *X* cannot know that *p*, because *p* is false. For a simple example, consider what happens when *X* sees someone over 215cm tall. She knows that he is tall, and even knows that she knows this. But the actual case is nearby to one where the person she sees is 214.5cm tall, which is nearby to one where he is 214cm tall, which is…, which is nearby to one where he is 140cm tall, where she clearly cannot know that he is tall, because he is not tall. So the KK principle, applied to the interpretation of box as 'knowably' plus the margin of error principle has led to the contradictory conclusion that we can know falsehoods. Something has to go, and Williamson says it is the KK principle. Hence, he thinks, the fact that his generation of higher-order vagueness relies on rejecting the KK principle is no reason to reject that generation.

Here's a different kind of problem for this account of higher order vagueness. Let us grant that we can be ignorant about whether we know that the salient guy is tall even if we actually know that he is tall. We noted above that not just any kind of ignorance generated the kind of indeterminacy distinctive of vagueness. The ignorance has to be of a particularly semantic kind. In particular, according to the above definition, it has to be ignorance of what a particular proposition *says*. But it is not so clear we *can* be ignorant of these matters in the crucial case. So just applying the above definition, we get

> *Definitely a is definitely tall* is true iff for all propositions *p*, if *X* cannot know that it is not the case that *a is definitely tall* says that *p*, then *p*

The idea seems to be that since we cannot know that *a is definitely tall* says that *a* is above 182.2 cm tall, say, since we only know that the definitely-tall/not-definitely-tall borderline falls between 179.5cm and 182.5cm, not precisely where it falls, *Definitely a is definitely tall* is false. The problem is that we can be certain that *a is definitely tall* does not *say* that *a* is above 182.2 cm tall. As per the above definition, what it says is:

> For all propositions *p*, if *X* cannot know that it is not the case that *a is tall* says that *p*, then *p*

Of course we don't know which propositions are such that *X* cannot know that it is not the case that *a is tall* says that *p*, and that is why we don't know that *a* is definitely tall. But since this isn't any ignorance about what is *said* by a sentence, this no more shows that we have higher-order vagueness than our ignorance about rabbit qualia shows that we have first-order vagueness in certain sentences about rabbit qualia. This looks like a technical problem, and it's a law that all technical problems have technical solutions, so this one has one too, though it's not as simple as you might think.

As a first attempt, note that we don't know whether the biconditional *a is definitely tall iff a is taller than 182.2cm* is true. Perhaps this is the right kind of semantic ignorance to generate vagueness. If it were, we would expect the following general definition of determinacy to work.

> *Definitely s* is true iff for all propositions *p*, if *X* cannot know that it is not the case that *s* is true iff *p*, then *p*

The idea, as you may have noticed, is to replace an intensional theory of meaning with an extensional theory of truth as the foundation of the definition of definiteness. And, as you may have also noticed, all the normal problems that arise when you send an extensional concept to do an intensional concept's work arise. Let *s* be a true but unknowable sentence about rabbit qualia, and consider what happens when *p* is an obviously false proposition, like 0=1. Since we do not know whether *s* is true, we don't know that it is not the case that *s* is true iff 0=1, but of course 0 does not equal 1. This means that *s* is not definitely true, whereas it should be definitely true. The following patch, which reinstates intensionality in the most obvious way, doesn't work either

> *Definitely s* is true iff for all propositions *p*, if *X* cannot know that it is not the case that necessarily *s* is true iff *p*, then *p*

The problem with this is that it is possible for *s* to mean something different to what it actually means. Indeed, *s* could mean anything at all, at least modulo constraints on what could possibly be meant by a sentence. (See the discussion of the modal paradoxes in *Plurality* for a sketchy but persuasive argument that some propositions can't be meant by any intentional entity, and hence can't be meant by sentences in a public language.) So for any proposition *p*, *X* can know that it is not the case that necessarily *s* is true iff *p*. So *Definitely s* is true for any *s* whatsoever. This won't do, but only a small change will get us to formulation that *looks like* it will work.

Of course it is possible that *s* means that *p* for any old *p*. It is less obvious that it is possible that *s* could mean this in English, since you might think that that if *s* meant something radically different, this would show that the language in question was no longer English. But English is flexible, and if *English* is not a rigid designator then it is more flexible still. Fortunately, it is less obvious again that *s* could mean *p* this in this language, where the demonstrative picks out the language currently being used. (I hope this is English, but maybe I'm perverse enough a user of words to be speaking a different language. As the saying goes, America and England are two countries separated by different languages, or something like that.) And since Kaplan showed that demonstratives are rigid designators, we do not have to worry about the possibility that English might have been radically different to the way it actually is, for we can be certain that this language could not have been radically different to the way it actually is. This suggests the following definition of definiteness.

> *Definitely s* is true iff for all propositions *p*, if *X* cannot know that it is not the case that necessarily *s* is true in this language iff *p*, then *p*

The scope of various terms there might be problematic, so here it is in symbols:

$$\mathrm{True}(\ulcorner \mathit{Def}\ s \urcorner, l) \leftrightarrow [\forall p\ (\neg K_X[(\neg\Box \mathrm{True}[s, l]) \leftrightarrow p)]) \rightarrow p]$$

The square and round brackets don't have any different meanings, they are just there to make it a little easier to track visually what the scope of every term in the sentence is. This definition seems to avoid all the problems. We are ignorant of whether certain necessitated biconditionals like *Necessarily ⌜a is definitely tall⌝ is true in English iff a is over 182.2cm tall* are true, so *Definitely a is definitely tall* do fail to be true. So this can all be formalised in a way consistent with the existence of higher-order vagueness, even if the formalisation is hideously ugly. As famous leader once said, ten out of ten for good thinking, but minus several million our of ten for style. I leave it to the reader to judge whether this still counts as an *elegant* solution to the problem of higher-order vagueness.

4.5. Epistemic and Doxastic Problems

So the epistemicist holds that there are all sorts of hidden boundaries around. The apparent vagueness of some terms is not due to the fact that they have no sharp boundaries, but because we cannot know where those boundaries are. 'Apparent' here might be thought to be a weasel word, because if epistemicists are right then the vagueness in various terms just does consist in their having boundaries that are unknowable in the right kind of way.

If there are these sharp boundaries, then there are a few puzzling questions to which the epistemicist owes us an answer. Most of these are variants on the following question: why does it *seem* that there are no such sharp boundaries? Three ways of sharpening this question come to mind, as listed here.

(1) Why is it that we don't, and apparently can't, know where these boundaries lie?
(2) Why is it that we don't, and apparently can't, have justified true beliefs about where these boundaries lie?
(3) Why is it that we don't even attempt to discover where these boundaries lie?

As noted already, Williamson has an impressive answer to (1). But it may not be clear how this is meant to translated into answers to (2), which has been pressed by Crispin Wright, or (3), which has been pressed (separately) by Rosanna Keefe and Hartry Field. Let us deal with these in order.

The core of the answer to (1) is that any knowledge about the location of the boundary would violate margin of error principles for knowledge. These principles say, roughly, that if you know *p*, then *p* must be true in all nearby cases. No such principle holds for justified beliefs. You can have a justified belief in *p* even if *p* is not true, so you can certainly have a justified belief in *p* even if *p* is not true in a nearby situation. So you'd expect that one could have a justified true belief in *p* even when *p* is false in nearby situations, or even in most nearby situations. Indeed, such cases are well known. They are normally called Gettier cases. So no explanation in terms of margins of error will help answer (2). Yet it seems (2) is a legitimate question to ask. Maybe this is a mistake, but it seems that we could not have a justified true belief about where one of these hidden boundaries lies. And it seems that this is something that stands in need of explanation, and the epistemic theory does not obviously explain it.

Here is one possible explanation that seems like it should be satisfactory to an epistemicist. (This was suggested to me by Juan Comesaña and Alyssa Ney, as was the answer to (3) discussed below.) Unlike knowledge, justification comes in degrees. What it is to not be justified in believing *p* is to not have a very high degree of justification for *p*, not necessarily to have no justification for believing *p* at all. Now there are a few things we know about how justification relates to knowledge. Let *p* and *q* be propositions such that you could never know which of them were true, were one of them true. Formally, the only way you can know $(p \vee q) \supset p$ or $(p \vee q) \supset q$ is by knowing that $\neg(p \vee q)$. Then it seems to follow that you couldn't be much more justified in believing *p* than in believing *q*, for it is if the reasons you have that give you more reason to believe *p* than *q* can *never* provide grounds for knowing $(p \vee q) \supset p$ without knowing $\neg(p \vee q)$, they cannot be particularly strong. Further, if there are many pairwise inconsistent propositions $p_1, p_2, \ldots, p_n$ such that for any two you are not much more justified in believing one than the other, your degree of justification for believing any of them is rather small. (Compare the equivalent claim for probabilities: if there are many pairwise inconsistent propositions $p_1, p_2, \ldots, p_n$ such that for any two the probability of one is not much more than the probability of the other, the probability of each is rather small. Since degrees of justification are not probabilities, the connection between these two principles is not immediate, but I think the principle about probabilities does provide a kind of initial plausibility to the principle about justification.) The explanation should now be straightforward. For any two hypotheses about where the borderline is, you cannot know either of the following conditionals: if one of these hypothesis is true, it is the first; if one of these hypotheses is true, it is the second. And the epistemicist has an explanation for this. Further, there are many such hypotheses, which are all pairwise inconsistent. Hence your degree of justification for any such hypothesis is rather low. That is, you are not justified in believing such a hypothesis.

Similarly, although (3) raises difficult questions for the epistemicist, they can be answered. Often, when we are ignorant about something, we try and remove the ignorance. This is not our reaction to vague terms. We do not, as a rule, try and find where the boundary between the tall and the not-tall lies, as we may do if we were ignorant of it in a normal way. It seems the best thing the epistemicist can say about why we don't try and repair this ignorance is that it would be impossible for us to do so. But this answer is doubly defective. First, we all try and do impossible things sometimes. (Think Hobbes trying to square the circle, if not the White Queen encouraging Alice to do six impossible things before breakfast.) Secondly, we only realise that this is impossible if we are epistemicists. And, as even epistemicists must admit, epistemicism is not the natural response to vagueness. So why don't we look for the boundary?

The best explanation here is disjunctive. Some people are epistemicists. They don't look for the boundary because they believe it is impossible to know where it is. Other people are not epistemicists. As a rule, they do not believe that such a boundary exists. There are some exceptions that we shall meet in chapter 7, but this is certainly true as a rule. They do not search for a boundary for the simple reason that they believe there is none there to be found. This explanation is not particularly unified, but if the phenomenon to be explained is not particularly unified, which arguably it is not, this is no bad thing.

4.6. Metaphysical Problems

There is a different way of stating the intuitive problem with epistemicism that does not seem to rely on appeal to any epistemic or doxastic concept. Epistemicism requires that there be facts about where the boundary between the tall and the not-tall lies, but intuitively there could be no such fact. As John Burgess puts it, epistemicism has no clear answer to the following question:

(Q) If vague concepts really do have sharp boundaries, what determines where those boundaries lie?

As Burgess notes, Williamson has had a few of attempts to answer this question, though none of them seem entirely successful. I will mostly follow Burgess's exegesis of Williamson here, except at one crucial point where I am sure Burgess gets the epistemic theory wrong. I think Burgess's mistake can be corrected, but I am fairly surely it is a mistake. Burgess starts by noting four kinds of answer the epistemicist can give, and that indeed Williamson has given, to (Q).

Austere: Provide answers that are unsatisfying, but are strictly speaking answers.
Indirect: Show that epistemicism can endorse various supervenience theses related to (Q), and suggest the truth of these thesis is sufficient to answer (Q).
Parasite: Wait for the anti-epistemicist, or as Burgess puts it the *indeterminist*, to answer (Q), or something like it that does not assume the existence of sharp boundaries, and show that epistemicism can endorse that answer.
Illegitimacy: Argue that the demand for an answer to (Q) is illegitimate, so it is no harm that epistemicism cannot provide an answer.

Here is an illustration of the austere strategy. Williamson, and Burgess, use *heap* rather than *tall* in their example, but the crucial points seem to be the same, and there are fewer extraneous complications when we use *tall*. Assume, as above, that any adult American male above 179cm tall is tall, anyone of them at or below that height is not tall. We can then put two questions to the epistemicist.

(4) What makes it the case that 179 is the threshold, rather than 179.1, or 178.9?
(5) Of two people, *a* and *b*, indistinguishable when viewed under optimal conditions, what makes *a* tall and *b* not tall?

There is a simple answer to (4): 179 is the threshold because people above that height are tall and people below that height are not. And there is a simple answer to (5): *a* is tall because he is above 179cm tall, and *b* is not tall because he is not. Obviously these claims are true (given the assumptions). And they have the *form* of answers. Indeed, Burgess points out that either one of them might even count as an explanatory answer, when taken on its own. But taken together, they are clearly not explanatory. And, intuitively, we had a right to an explanatory answer to the two questions.

The second move is to investigate why we might have believed that (Q) had explanatory answers. One reason is that we might have thought that meaning had to be derivable from use. Formally, we might spell this out in one of the following three ways.

(S1a) Meaning supervenes on use.
(S1b) Meaning is knowable on the basis of knowing use.
(S2a) Vague truths are supervenient on precise truths.
(S2b) Vague truths are knowable on the basis of knowing precise truths.

Williamson quickly points out that epistemicism agrees with the two a-theses. According to the epistemicist, the meaning of a word could not be different unless the use of that word was different. (We understand 'use' here in a broad enough way to include the circumstances in which the term is used, so (S1a) is compatible with all sorts of varieties of semantic externalism, even if they are mostly false.) What he denies are the two b-theses. We could know all there is to know about use, and we could know all the precise facts there are, and still not be in a position to know whether some particular person is tall. I think, and this is a little speculative because I don't think the crucial texts are particularly perspicuous here, the explanation of why we think (Q) has an explanatory answer is that we think the two b-theses are true. Here's how one could argue from the b-theses to the claim that (Q) has an explanatory answer. If meaning is knowable on the basis of knowing use, then we must already know, implicitly, the broad outlines of the function from meaning to use. Any way of making that implicit knowledge explicit would constitute an answer to (Q). But we can always make implicit knowledge explicit, so there is an answer to (Q). And since the b-theses are false, this kind of reasoning is unsound, even if it is attractive.

Williamson's reasons for rejecting the b-theses, and hence for rejecting this kind of reasoning, are a little slim. He says that (S2b) commits us to "a form of scientism, on which all questions can be replaced without significant loss by questions of natural science." Well what's wrong with that! Of course as phrased this form of scientism is probably false, since questions concerning self-location probably cannot be replaced without significant loss by questions of natural science. See the amusing passage from "Attitudes *De Dicto* and *De Se*" where Lewis bemoans the fact that his theory makes him *sound* like an anti-scientific subjectivist. But as Lewis notes, this is really the only exception to the rule that science gives us a *complete* picture of the world.

As Burgess notes, the a-theses do not really provide all we want in a theory of meaning, and in fact they fail in a way that seems particularly relevant to the kind of objection to epistemicism that we have been considering. The truth of the a-theses is compatible with the function from use to meaning, or from precise truths to vague truths, being completely unsystematic. But we intuit that meaning is not just determined by use, it is determined by use in a systematic way. Here is how Burgess puts the point.

> If *n* marks the boundary between heaps and non-heaps, we feel strongly inclined to say that locating the boundary precisely here is *arbitrary*. But how can a boundary be arbitrary when there is no arbiter? Clearly we speakers have not arbitrated and there

> seem to be no facts about our behaviour (collectively) which simulate for us an arbitration…[It] is implausible to hold that this behaviour is sufficiently refined to select sharp (bivalent) boundaries. (514)

This seems to me to be exactly right, though there is still one turn to go. We have not yet talked about the parasite strategy. Recall Sider's argument for indeterminacy of meaning. It said that terms must be indeterminate because there were different candidate meanings that did not differ with respect to how well they fit use or with respect to how eligible they were to be meanings. What we have discussed so far has been attempts to find relationships between meaning and use that allow the epistemicist to answer (Q) or argue that our intuition it has an answer is misplaced. The parasite strategy appeals less to facts about use and more to facts about eligibility. (The following is quite removed from Williamson's *presentation* of the view, but I think is a fair and accurate translation of his ideas into the conceptual framework we've been working in thus far.)

The indeterminist does not hold that meanings of vague terms are *completely* indeterminate. To the extent that they think (Q), or something like it that does not assume precision, has an answer, they must think there is something explanatory to say about the connection between use, eligibility and meaning. The epistemicist can adopt that answer and just append to it a slightly more detailed story about eligibility. Williamson holds, in effect, that more restrictive meanings are *ceteris paribus* more eligible than less restrictive meanings. When faced with a choice between competing concepts which each have a claim to be the meaning of *t*, the meaning function selects the one satisfied by the fewest elements. As Williamson puts it, truth and falsehood are not symmetric. If something fails to meet the requirements for being true, then it is not true. The arbitrariness is resolved in virtue of stinginess.

This assumes that there is a most restrictive possible meaning. In theory this looks like a ridiculous assumption. Juan Comesaña pointed out to me that it won't work too well if Williamson's indeterminist opponent has a theory of content on which there is higher-order vagueness. Given the difficulties that indeterminist theories have had so far with higher-order vagueness, I am prepared to ignore this difficulty. I don't think I am begging questions against Williamson in doing so. There are cases in which there is a determinate set of candidate meanings, but no 'most restrictive meaning', which raises problems for the asymmetry theory. If the candidate meanings for our term are properties like *More than 179cm tall*; *More than 180cm tall*; etc. then there will be a most restrictive candidate in the sense that there will be a concept that is satisfied by all the objects that satisfy any of the candidate concepts. It is not so clear this will be the case with predicates generally. Consider, for example, the predicate *talented artist*. This has many candidate meanings, but is there a 'smallest' meaning that includes all the things that satisfy any of the candidates? Perhaps, but it is not so clear. Burgess has a (bad) reason for thinking that this theory of Williamson's cannot work for predicates.

> Consider a borderline case [of a colour patch] that fails to pass the test for being red and also fails to pass the test for being orange. [I.e. it is not red on the most restrictive meaning for 'red', and is also not orange on the most restrictive meaning for 'orange This will occur if there is any abstention or disagreement at the border. Given asymmetry, it is neither red nor orange, for it passes the test for being not red and also the test for being not orange. But, on the epistemicist view, there is a sharp boundary in the series between red and orange; every patch is either one or the other. (519)

The last line here is a mistake. It is *not* part of the epistemicist theory that there is a sharp boundary between the red and orange patches. It *is* part of the theory that there is a sharp boundary between the red patches and the not-red patches, but this does not imply that there is a sharp boundary between red and orange. It would imply this if we had some principle like: anything in this region that is not red is orange, but I see no reason to adopt such a principle.

If such a principle were part of the meaning of *orange*, Williamson could easily adjust his position to avoid the problem. Many concepts are atomic - their meaning is not a logical construction out of other concepts. The asymmetry theory of meaning sketched above applies to them. Other concepts are not atomic. If *orange* has as part of its meaning *not red*, then it will not be atomic. (Perhaps *orange* just means *a colour between red and yellow that is neither red nor yellow*). The meaning of these is generated by the meaning of their atomic parts, so the asymmetry theory does not directly apply. If this is the meaning of *orange*, which I very much doubt, all the borderline cases will be orange.

So *orange* and *red* do not pose a problem for Williamson. There is a real problem for Williamson's approach, though, when we focus on the kinds of cases prevalent in discussions of the Problem of the Many. The following is basically David Lewis's example from "Many, but
Tibbles is a cat, and he is shedding hair. Some of his hairs are fairly loosely attached to him, so loosely in fact that you might think that they have ceased to be parts of Tibbles. That is, for some hairs, it is indeterminate whether they are part of Tibbles. Assuming Tibbles to be not near any other cats, these hairs are not part of any other cat, so they are either part of Tibbles, or part of no cat. The upshot of this is that Tibbles has some parts that are determinately parts of him, and may have some parts, mostly hairs, that are indeterminately parts of him. Call the fusion of the parts that are determinately part of Tibbles Tib_0, and each fusion of Tibbles with some of the indeterminately attached parts Tib_1, Tib_2, …, Tib_n. Now consider each of the following sentences:

(T0) Tib_0 is a cat.
(T1) Tib_1 is a cat.
…
(T*n*) Tib_n is a cat.

Intuitively, one of these is true, since for some *j*, Tibbles is Tib_j, and Tibbles is a cat. (If you don't think Tibbles is identical with the fusion of his parts, perhaps because he has a different modal profile, replace *is a cat* in this discussion with *is exactly co-located with a cat*. I prefer the simpler formulation, but if it is metaphysically objectionable, I don't need to rely on it.) But none of the Tib_j "pass the test" for being a cat, since on the indeterminist's theory, for each Tib_j there is a meaning for *cat* according to which it is not a cat. Hence each sentence (T*j*) is false, according to the asymmetry theory. Since Tibbles is Tib_j for some *j*, this implies that Tibbles is not a cat. We're back to Peter Unger's version of nihilism! Well, it might be plausible to say that there are reddish, orangish patches that are neither red nor orange. But it is not particularly plausible to deny that Tibbles is a cat, so Williamson is in some difficulty here. The important point is that there might be candidate extensions E_1 and E_2 for some term, say *cat*, such that we do not know, and cannot know, which of them, if either, is the correct one, but such that we do know that their intersection is not the extension of the relevant term. This means that there is no 'most
cat, and Williamson's appeal to it is doomed to failure.

Similar points can be made once we move the discussion away from predicates and towards names. Hartry Field notes that the following example, originally due to Robert Brandom, is troubling for the epistemicist in this context. Brandom's example is somewhat artificial, but unless artificial cases can be ignored, it seems to pose an insurmountable problem for the epistemicist. In our language the names for the two square roots of -1 are *i* and -*i*. Imagine a linguistic community that has two atomic names for these two roots, rather than our one atomic name (*i*) and one compound name (-*i*). Brandom suggests that they name the roots / and \. The community knows that / = -\, but they see that as no reason to drop the name / for that root of -1, in favour of the longer symbol -\. After all, they note, it is also the case that \ = -/, so there is no more reason to drop the symbol / than to drop \.

It seems possible that there could be such a community, and possible we could come to learn their language. Perhaps this is more of a stretch, but it seems possible that we could come to start speaking a hybrid language, where we incorporate some of their words into our native language. This kind of hybridisation happens all the time in the real world, so it does not seem outlandish to imagine it happening here. So, by assumption, we have started speaking a language where as well as *i* and -*i* as the names for square roots of -1, we have / and \. Now consider the sentences (8) and (9).

(8) / = *i*
(9) \ = *i*

By the asymmetry theory, these are both false, since neither of them "passes the test" for being true. Of course, if / just meant -\, or \ just meant -/, we could say that the asymmetry theory does not apply. But this is not how the case is developed. Both / and \ are primitive names, so if the asymmetry theory applies it applies to all of them.

The problem is that it is impossible that both (8) and (9) be false. There are only two square roots of -1, so if / and \ are both square roots of -1, and neither equals *i*, they must both equal -*i*. But this implies they equal each other. And this implies that /-\ = 0, which in turn implies that $(/-\backslash)^4/16 = 0$. But it is easily provable within the original community's mathematical theory that $(/-\backslash)^4/16 = 1$. So the asymmetry theory has the unfortunate consequence that 0=1. Hence it is false.

In summary then, Burgess's argument looks shaky as he applies it, but when we study more cases looks like it might have some bite. So Williamson needs to say more to make the parasite strategy work, and it is not so clear where he can look.

4.7. Sorites

I said above that I didn't know what the supervaluational solution to the Sorites was meant to be. Well, I also do not know what the epistemicist solution to the Sorites is meant to be. As they say in football, the following might be a hidden indicator statistic. In *Vagueness* there are 61 index references to the Sorites paradox. Exactly six of these appear in the chapters where Williamson is setting out his positive theory. Four of these occur in the argument against the KK principle noted above. The other two occur in an argument for the existence of hidden boundaries in meaning that turns on the behaviour of omniscient speakers. And that's it! There's no positive discussion of the paradox, or how the epistemicist proposes to solve it.

Of course, we all know where the epistemicist thinks the Sorites paradox goes wrong. In a series of premises version of the Sorites, exactly one of the premises is false, and in a quantified version, the quantified premise is false. But that no more solves the paradox than a bare statement that this is the best of all possible worlds solves the paradox of evil. A solution must explain why we ever thought the premises were true. And I doubt the epistemicist is in a position to explain this, at least without supplementing the theory in some way. (Delia Graff makes a similar point to this, though with less reliance on hidden statistical indicators like index references.) The problem is that the epistemicist explanation of why there *appears* to be truth value gaps prevents a natural explanation of why the Sorites premises seem to be true. Indeed, this explanation of the apparent truth value gaps creates a distinctive problem for attempting to explain the apparent truth of Sorites premises.

As previously noted, Williamson argues that we should only be prepared to assert sentences that we know. This explains why, in cases of vagueness, we are not prepared to assert that the subject has the property in question, or that she does not. Return to Louis, our borderline case of baldness. If there is a sharp boundary between the bald and the not bald, and we do not know where it is, then we do not know whether Louis is bald, for we do not know what side of the boundary Louis falls on, no matter how much we know about Louis's intrinsic properties. Hence we should not be prepared to assert that he is bald, nor to assert that he is not bald. These predictions seem to be borne out in practice. If speakers in general confuse this unwillingness to assert a sentence, even in ideal circumstances, with the sentence failing to be true, we have an explanation for why there seem to be truth value gaps.

The problem is that our ignorance of where the boundary lies also means we are ignorant about whether several Sorites premises are true. If, for all we know, having 150 hairs is the boundary between being bald and being not bald, then we do not know that if a person with 150 hairs is bald, then so is a person with 151 hairs. Hence we should not be prepared to assert the conditional: If a person with 150 hairs is bald, so is a person with 151 hairs. Nor should we be prepared to assert the negated conjunction: It is not the case that a person with 150 hairs is bald and that a person with 151 hairs is not bald. Yet the whole problem concerning the Sorites is that we are prepared to assert these things.

The difficulty here is not just that one of the Sorites premises is false, and hence we cannot know it, and hence we should not be prepared to assert it. (Though that alone is a problem.) The real problem is that for *every* Sorites premise concerning borderline cases, we do not know that premise is true, and hence we should not be prepared to assert it. Just like the supervaluationist, the epistemicist seems to predict that we will demur from all the crucial premises in a Sorites argument. In a version of the Sorites that uses negated conjunctions, the only Sorites premises that should seem acceptable where we can tell which of the conjuncts is false. In all other cases, we do not know that the borderline falls just 'between' the two conjuncts, and hence that both are true. But what gives the Sorites its interest is that we are prepared to accept all the negated conjunctions as premises, even though in many cases we cannot say just which of the conjuncts is false.

This does not show that epistemicism is false, though it does show that the theory is incomplete in a way that might have seemed fairly important. In chapter 8 I will consider a theory of the pragmatics of compound sentences that *might* get the epistemicist out of this trap, though I will conclude that the theory does not fit as well with epistemicism as with theories on which vagueness is a semantic phenomenon. The important point to note is that epistemicism needs an explanation for the attractiveness of the Sorites, and right now it doesn't have one.

4.8. Transparency and Translucency

Mario Gomez-Torrente and Delia Graff have argued that there seems to be a problem with incorporating higher-order vagueness into Williamson's preferred semantics for knowledge claims. The problem starts with some observations about the sentence (10).

(10) A man with no hair whatsoever on his head is bald.

Gomez-Torrente and Graff claim, and Williamson accepts, that (10) is true, and it is determinately true, and it is determinately determinately true, and so on. Relatedly, they claim that (10) can be known to be true, and it can be known to be known to be true, and it can be known to be known to be true, and so on. The point is that even if *bald* is vague, there can still be absolutely clear cases of its application.

As an aside, I am not so sure that (10) is as obvious as they claim. Some people argue that a man who has no hair on his head because he has shaved it all is not bald. They are wrong, such a person is bald. I'm quite confident of this, the semantics for *bald* does not discriminate between causes of hairlessness. Am I confident that I know I know the semantics for *bald* is so indiscriminate? At about this level I start to have my doubts. So maybe (10) is not such that one can know that one knows that one knows that … that one knows it, no matter how often one iterates the 'knows that'. But even if it is not, it seems that some modification of (10) should have this property. Perhaps we could replace (10) with *A man with no hair whatsoever on his head who has never had any of his hair removed by artificial means is bald*. Or perhaps we could pretend that (10) makes a claim sufficiently precise that such concerns are addressed. Not only can we, we will! End of aside.

Williamson provides the following semantic models for knowledge claims. (We will return to the question of whether the operator K in these models should be interpreted as *It is known that* or *It is knowable that*.) A model is a four-tuple of a set of (uninterpreted) worlds, a valuation function that assigns sets of worlds to atomic proposition, a distance function d that takes (unordered) pairs of worlds as inputs and provides a real number as output, and a value r that is the permissible 'margin of error' of knowledge. Williamson then provides two accounts for what it is for p to be known in a world.

Fixed Margin semantics: KA is true at x iff for all worlds y such that $d(x, y) \leq r$, A is true at y.
Variable Margin semantics: KA is true at x iff there exists a positive real value ε, such that for all worlds y such that $d(x, y) \leq r + \varepsilon$, A is true at y.

The fixed margin semantics says that for KA to be true, A must be true at all worlds within r of the world in question. The variable margin semantics says that for KA to be true, there must be some sphere of worlds, of radius greater than A, such that A is true throughout the sphere. This might seem like a very small change, but the effect of it is that $A \rightarrow \mathrm{K}\neg\mathrm{K}\neg A$ ceases to be true in all models. This is true in all fixed margin models, but at times it seems implausible. The mere truth of A should not be enough to make it the case that we can know that we cannot know that $\neg A$. If A is about some subject on which we are necessarily deeply ignorant, perhaps this is implausible when K is interpreted as *It is knowable that*. It is certainly implausible when K is interpreted as *It is known that*. Because our interest here is in the connection between knowledge and vagueness, we are more interested in the case where it is interpreted as a knowability operator. This will lead to problems enough.

For simplicity, we will write $B(n)$ as short for ⌜Any man with exactly n hairs on his head is bald⌝. Following Graff, we will say that a person with n hairs is transparently bald iff $B(n)$ is known, and it is known that $B(n)$ is known, and so on, and that a person with n hairs is translucently bald iff it $B(n)$ is knowable, and it is knowable that $B(n)$ is knowable, and so on. We will always interpret K as *It is knowable that*, and K* as *It is translucent that*. Gomez-Torrente claims that $\mathrm{K}^*B(0)$, and there is a value z such that $\neg\mathrm{K}^*B(z)$. Since there are values z such that $\neg B(z)$, this last claim is not too implausible. If epistemicism is true, then there is a boundary between those values x for which $\mathrm{K}^*B(x)$ is true, and those values for which it is not true. That is, for any value ε, there are values v and w such that $|v - w| < \varepsilon$, and $\mathrm{K}^*B(v) \wedge \neg\, \mathrm{K}^*B(w)$. This of course is no objection to epistemicism, just a statement of the epistemicist position. The problem is that it follows from the margin of error semantics plus a very small assumption that it is knowable where this boundary falls. And if this is true, then according to the epistemicist, then the predicate *translucently bald* is not vague.

Gomez-Torrente's original argument for this result did not quite cover all the cases, so we will concentrate on Graff's version of the argument. She proves that if the model is a 'stepping-stone' model, then it is knowable where the boundary between the translucently bald and the not translucently bald lies. A model is a stepping-stone model iff for any two worlds a and b such that $d(a, b) \leq 2r$, then there is a world u, the stepping-stone, such that $d(a, u) \leq r$ and $d(u, b) \leq r$. To motivate the idea that there are such stepping stones, just note that there are no 'discontinuities' in modal space. The reason that we can generate Sorites paradoxes for any predicate at all is that we can find a continuous series of possibilia between a clear case of an *F* and a clear case of a non-*F*. Just as we can do this for possibilia, we can do it for entire possibilities, which is all that the stepping stone assumption says.

(Actually, there is one sense in which Graff's assumption is too strong. It says that whenever the distance between a and b is exactly $2r$, then there is a stepping-stone *exactly* half-way between them. A mere appeal to the plenitude of possibilities does not seem to support this claim. But Graff does not need anything this strong. Her argument goes through on the weaker assumption that there is some value $c > 1$ such that $d(a, b) \leq cr$, then there is a world u, such that $d(a, u) \leq r$ and $d(u, b) \leq r$. This can be verified by just inspecting where she uses the stepping-stone assumption in the proof. And this assumption does seem to be justified by the fact that possibilities are so plentiful.)

How bad a result is this? As Graff points out, *translucently bald* is not a commonplace enough predicate that our intuitions about it are guaranteed to be particularly sound. So if, counterintuitively, it turns out to not be vague, this is not a particularly large concern. Further, there are some simple changes the epistemicist can make to avoid this result. If we hold, for instance, that r is not a constant, but a function from worlds to values, then the proof that *translucently bald* is not vague does not go through. This technical change would amount to the somewhat odd assumption that how large the margin of error for knowledge is depends on the nature of the world. If this move seems strange, note that if we deny it, i.e. accept that r is a constant, and accept that for any two worlds $d(a, b)$ is a real number, then only claims that are true in all worlds will be translucent. But the latter result is implausible, if we think that a person with no hair is translucently bald. The problem is not that our model should, intuitively, have worlds in which a person with no hair is not bald. Obviously, it should not. Rather, our model should have worlds in which the word 'bald' does not apply to people with no hair. We know those worlds are non-actual, but we cannot rule them out *a priori*. So there are a couple of ways in which the epistemicist can respond to this particular objection, but they require either giving up some intuitively plausible claims, or adjusting the semantics Williamson offers.