GARY EBBS

VAGUENESS, SHARP BOUNDARIES, AND SUPERVENIENCE CONDITIONS

In his impressive book *Vagueness*, Timothy Williamson critically surveys the entire literature on vagueness and presents a brilliant new version of the theory that our vague concepts have unknown sharp boundaries.[1] His probing criticisms of previous views of vagueness are unified by a deep commitment to realism and a correspondingly thorough rejection of definitional theories of meaning and consensus theories of truth. I share Williamson's commitment to realism, and I find many of his arguments persuasive. As I see it, however, Williamson's proposed explanation of our ignorance of borderline vague truths faces a dilemma: either we have no grounds for accepting it, or it is no more than an elaborate *restatement* of what it is supposed to explain – our starting observation that we do not know borderline vague truths. This dilemma discredits Williamson's underlying methodological assumption that we can distinguish between "conceptual" and "empirical" sources of our ignorance of borderline vague truths.

To show why, I will focus on an intuitively plausible assumption that many philosophers simply take for granted. The assumption is best seen as a generalization from intuitions about the conditions for applying particular vague concepts. For instance, consider the intuition that *if any two individuals x and y have the same number of hairs on their heads*, *then x is bald if and only if y is bald.* If we take this for granted, then if we know, for example, that Albert is bald and Bob is not bald, we will infer that Albert and Bob do not have the same number of hairs on their heads. If we also know that Albert has 10 hairs on his head, we will infer that *every* individual with 10 hairs on his head is bald. More generally, the intuition about baldness implies that *an individual with n hairs on his head is bald only if every individual with n hairs on his head is bald.* Both the intuition that *if any two individuals x and y have the same number of hairs on their heads*, *then x is bald if and only if y is bald*, and the consequence that *an individual with n hairs on his head is bald only if every individual with n hairs on his head is bald* are what I call substantive supervenience conditions for being bald. I call them "substantive" because they generalize

about relations between an individual's being bald and the number of hairs on his head, and thereby tell us more than that to be bald one must be bald. For a wide range of different vague concepts, there are intuitively plausible substantive supervenience conditions such as the two just stated for the concept bald. Generalizing from substantive supervenience conditions that they find plausible, many philosophers assume that for *every* vague concept there are substantive supervenience conditions, whether or not we know what those conditions are.

For some accounts of vagueness, such as the supervaluation approach that I will discuss below, the intuitive plausibility of this assumption is a good reason for treating it as a constraint on our understanding of vagueness. If we accept Williamson's account of vagueness, however, there are three considerations that taken together should lead us to doubt the assumption. First, bivalence for vague terms does not by itself guarantee that there are substantive supervenience conditions for vague concepts, so Williamson's argument for bivalence, even if it were completely successful, would not by itself support that assumption; second, since Williamson rejects the supervaluation approach, he cannot appeal to meaning postulates or constraints on valuations to vindicate the assumption; and third, Williamson's own realistic view of the relation between meaning and use should lead us to doubt that there are substantive supervenience conditions for vague concepts.

Williamson's explanation of our ignorance of borderline truths does not officially rest on the assumption that we know substantive supervenience conditions for vague concepts. But I will argue that if we do not know any such conditions, then either we have no grounds for accepting Williamson's explanation of our ignorance of borderline truths, or his explanation is just an elaborate restatement of what it is supposed to explain. The methodological moral of this dilemma is that once we give up the supervaluation approach to vagueness, we must also give up the initially attractive idea that we can distinguish between "conceptual" and "empirical" sources of our ignorance of borderline vague truths. I will present these points in order, outlining Williamson's arguments and theory along the way.

1. THE PROBLEM OF VAGUENESS AND A BRIEF SKETCH OF WILLIAMSON'S SOLUTION

According to Williamson, "the matter of vagueness gets its urgency from sorites paradoxes" (p. 72). A sorites paradox for 'bald' may be formulated as follows:

(1a)   A person with 0 hairs on his head is bald,

(2a) $\forall n$[(a person with n hairs on his head is bald) $\rightarrow$ (a person with $n + 1$ on his head is bald)],

(3a)   A person with 100,000 hairs on his head is bald.

If we use 'P( )' to stand in for 'a person with ( ) hairs on his head is bald' we can see this argument as an instance of the following form:

(1)   P(0),

(2) $\forall n$[P(n) $\rightarrow$ P($n + 1$)],

(3)   P(100,000).

The argument is paradoxical because the conclusion seems clearly false, yet both premises seem true. We feel confident that anyone with 0 hairs on his head is bald, and we seem forced to accept the second premise, since we feel sure that 1 hair cannot make the difference between being bald and being not bald.[2]

One standard reaction to the sorites paradox is to question bivalence – the thesis that every statement is either true or false. Since for any vague term T, there are situations in which we would hesitate or even refuse to apply either T or T's negation, it is tempting to conclude that vague terms just do not have sharp boundaries. Some philosophers are so convinced of this that they *define* vagueness as the lack of sharp boundaries, and conclude that (2a), for example, is neither true nor false.

In contrast, Williamson's explanation of why sorites arguments are not sound is a version of the epistemic theory of vagueness, according to which bivalence holds for all vague utterances that "say something", and vague concepts have sharp boundaries. On this view, the second premises of sorites arguments, including (2a), are false; for each sorites argument, $\exists n$[P(n) $\wedge \neg$ (P($n + 1$)], even if we cannot know which number this is.

To identify the phenomenon of vagueness without begging the question of how to characterize it, Williamson gives examples. He assumes we can

agree that such terms as 'bald', 'heap', 'thin', and 'old' are vague without agreeing about what vagueness consists in (p. 2).

What is distinctive and ingenious about Williamson's version of the epistemic theory is his explanation of our ignorance of the hypothesized sharp boundaries of vague concepts. The basic idea is that our vague judgments amount to knowledge only if they are reliable, and in borderline cases they are unreliable. I'll examine Williamson's theory after I distinguish two senses in which a concept may be said to have sharp boundaries, and explain why Williamson rejects the supervaluation approach.

2. BIVALENCE, SHARP BOUNDARIES, AND INTUITIONS ABOUT MEANING

It is natural to think that if all vague utterances that "say something" are either true or false, then vague concepts have sharp boundaries. To evaluate Williamson's theory, however, it is crucial to distinguish between a weak and a strong sense of the phrase "concepts have sharp boundaries".

A concept has sharp boundaries in the weak sense just in case for all x, the term that expresses the concept is true of x or false of x. The law of excluded middle entails that for all x, x is bald or x is not bald; if all sentences in which the word 'bald' occurs are bivalent, then the concept bald has sharp boundaries in the weak sense.

It is tempting to think that if a concept has sharp boundaries in the weak sense, it also has sharp boundaries in the strong sense that there are substantive supervenience conditions for applying the concept. For example, if we suppose that the concept bald has sharp boundaries in the weak sense, we are inclined to conclude that there is a number n such that anyone with n hairs on his head is bald, and anyone with $n + 1$ hairs on his head is not bald.

We are inclined to conflate the weak and strong senses in which concepts may have sharp boundaries. If we accept bivalence and reject (2), we will accept

(4) $\forall n[\mathrm{P}(n) \wedge \neg(\mathrm{P}(n + 1)]$.

For instance, to accept bivalence and reject (2a) is to accept

(5) $\exists n[$(every person with n hairs on his head is bald) $\wedge \neg$(every person with $n + 1$ hairs on his head is bald)].

We are then inclined to conclude that

(6) $\exists n[$(every person with n hairs on his head is bald) $\wedge$ (every person with $n + 1$ hairs on his head is not bald)].

But (6) does not follow merely from the rejection of (2a). It is instructive to see why.

The second conjunct contained in (5) amounts to 'it is not the case that every person with $n + 1$ hairs on his head is bald', which, of course, is equivalent to

(7)   Some person with $n + 1$ hairs on his head is not bald

not to

(8)   Every person with $n + 1$ hairs on his head is not bald.

This means that (5), which amounts to the negation of (2a), could be true even if there are persons with $n + 1$ hairs on their heads who are bald. It does not follow from bivalence and the negation of (2a) that the boundaries of the concept bald are determined by the number of hairs on a person's head, so bivalence and the negation of (2a) do not imply that the concept bald has sharp boundaries in the strong sense. More generally, bivalence for vague terms does not by itself imply that vague concepts have sharp boundaries in the strong sense.

We are inclined to equate the falsity of (2a) with existence of sharp boundaries in the strong sense partly because we tacitly accept

(9)   For all x and y, if x and y have the same number of hairs on their heads, then x is bald if and only if y is bald.

(7) and (9) together imply (8),[3] so (5) and (9) together imply (6). But (9) is not a logical truth. We accept (9) because it captures intuitions about meaning.

These intuitions are also responsible for our unreflective acceptance of (1a). If we take (9) for granted, then to verify (1a) it is enough to find one person with no hair who is bald. If we don't take (9) for granted, however, then (1a) has the vulnerability of a universal generalization that is not logically true.

Our intuitions about meaning are what lead us to move so quickly from the conclusion that vague concepts have sharp boundaries in the weak sense to the conclusion that vague concepts have sharp boundaries in the strong sense. But what is the proper epistemological role for these intuitions?

3. MEANING POSTULATES AND SUPERVALUATIONS

One attractive answer, initially proposed by Carnap, is that we can choose to adopt meaning postulates that incorporate our intuitions about meaning into the "logic" of a semantically regimented language-system.[4] In particular, we can treat (9) as a meaning postulate. If, as we believe, there is someone with 0 hairs on his head who is bald and someone with 100,000 hairs on his head who is not bald, then (9), treated as a meaning postulate, guarantees that having 0 hairs on one's head is sufficient for being bald, and having 100,000 hairs on one's head is sufficient for being not bald. Given this meaning postulate, the property of being bald *supervenes* on the number of hairs on a person's head: two individuals cannot differ with respect to baldness unless they differ with respect to the number of hairs on their heads.

But Carnap's approach does not accommodate the intuition that many vague sentences are neither determinately true nor determinately false. To reconcile the spirit of Carnap's approach with this intuition, we need to capture what Kit Fine calls penumbral connections between sentences that contain vague terms, without assuming that all these sentences are either determinately true or determinately false.[5] In effect, Fine proposes that we treat a sentence such as (9) as a constraint on how the predicate 'bald' can be made more precise. On this supervaluation approach, (9) is a constraint on every admissible precisification, so there is a sense in which it is correct: it is true for all admissible specifications, or super-true. And even if we do not accept that there is some n such that for all admissible specifications ((every person with n hairs on his head is bald) $\wedge$ (every person with $n + 1$ hairs on his head is not bald)), (9) guarantees that for all admissible specifications there is some n such that ((every person with n hairs on his head is bald) $\wedge$ (every person with $n + 1$ hairs on his head is not bald)), so (2a) is super-false.

Like Carnap's meaning postulates, the supervaluationist's constraints on admissible specifications of the extensions of vague terms are theoretical explications of our intuitions about meaning. The intuition that the property of being bald supervenes on the number of hairs on a person's head is built into the supervaluation approach as a constraint on admissible specifications of the extension of 'bald'.

Williamson's central objection to the supervaluation view is that it cannot adequately explain higher-order vagueness. The objection begins with the observation that "the inadmissibility of a valuation is itself a vague notion" (p. 157). Williamson argues, in effect, that the supervaluationist faces a dilemma. On the one hand, we can eliminate higher-order vague-

ness only if we use a precise metalanguage to specify the meanings of vague sentences; but to use a precise metalanguage to specify the meanings of vague sentences would be to abandon "the central task of genuine semantics: saying what utterances of the object-language mean" (p. 191). On the other hand, if we cannot eliminate higher-order vagueness, then we cannot know which valuations are inadmissible, so super-truth is no better defined for vague sentences than truth (pp. 162–164).

In my view, the best reply to this argument would be to accept that higher-order vagueness can be eliminated only if we use a precise metalanguage, but deny that to use a precise metalanguage is to abandon the central task of semantics. The supervaluationist should claim that all languages have built in meaning postulates and constraints on admissible valuations, and the central task of empirical semantics is to say which language is used by a given population. Higher-order vagueness then amounts to indeterminacy in the phrase 'population P uses language L', not to vagueness in the meanings of expressions of any particular language.[6]

Unfortunately, to accept this reply on behalf of the supervaluation view, one must assume that speakers of a natural language such as English can be described without distortion as speakers of a precise language with explicit semantical rules and constraints on admissible valuations (or a "blur" of such languages, as David Lewis suggests).[7] In my view, this assumption is undermined by Quine's and Putnam's arguments against the logical positivists' analytic-synthetic distinction. Quine argues that the idea of a semantical rule of language is hopelessly unclear, and so he rejects the positivists' idea that some sentences of a natural language are analytic – true "in virtue of" semantical rules – and others are not.[8] Putnam argues that we have no criterion for judging whether or not a given natural language sentence that we currently hold immune to revision is true "in virtue of" semantical rules.[9] Quine and Putnam both conclude that natural languages do not have built in semantical rules, and so they would both reject the supervaluationist's assumption that speakers of a natural language can be described without distortion as speakers of a (blur of) precise language with explicit semantical rules and constraints on admissible valuations.

Suppose we agree with Quine and Putnam that natural languages do not have built in semantical rules. How then shall we understand what Williamson calls the central task of semantics – "saying what utterances of the object-language mean"? He seems to think that if we are allowed to use a vague metalanguage, it is not difficult to say what utterances of a vague object language mean, because the vagueness of an object-language word can be exactly matched by some vague word used in the metalanguage. To

test this idea, suppose we try to state in French the meaning of utterances of the English sentence 'Jones is bald', as follows:

(*)   'Jones is bald' veut dire que Jones est chauve.

The problem is that if English and French do not have built in semantical rules, there is no informative criterion for deciding whether the vagueness of 'bald' is *exactly matched* by the vagueness of the French word 'chauve'. How then can we determine whether (*) says what utterances of 'Jones is bald' mean? We can avoid this problem and still say what utterances of 'Jones is bald' mean only if we *use* 'Jones is bald' to say what utterances of 'Jones is bald' mean.

To summarize my argument in the last three paragraphs: the super-valuationist has a good reply to Williamson's claim that to use a precise metalanguage to specify the meanings of vague sentences would be to abandon "the central task of genuine semantics: saying what utterances of the object-language mean" (p. 191). The reply is that the central task of empirical semantics is to say which precise language is used by a given population. In my view, the only convincing way to challenge this reply is to use Quine's or Putnam's arguments against the assumption that natural languages have built in semantical rules. But Williamson can use those arguments only if he also accepts that there is no informative general criterion for deciding whether we have said what utterances of an object-language mean. This is compatible with (a deflationary interpretation of) his claim that the central task of semantics is to say what utterances of an object language mean, since even without an informative general criterion for accomplishing this task, we can simply *use* the sentences of a vague object language L to say what utterances of L mean.

I will discuss the consequences of Williamson's rejection of the super-valuation approach in more detail below. For present purposes, the crucial point is that once Williamson rejects the supervaluation approach, he cannot appeal to meaning postulates or constraints on valuations to vindicate the intuition that there are substantive supervenience conditions for vague concepts.

4. WILLIAMSON'S ARGUMENT THAT THERE CAN BE NO COUNTER-EXAMPLES TO BIVALENCE

Williamson's discussion of higher-order vagueness is meant to show that to give the meanings of expressions of a vague object language, we must use a vague metalanguage. His positive case for the epistemic view of vagueness

begins with his argument that if we use a vague metalanguage to give the meanings of expressions of a vague object language, the supposition that there are counterexamples to bivalence for vague sentences of the object language leads to contradiction.

The heart of Williamson's argument is that any adequate account of truth must accept the following biconditionals (p. 188):

(T)   If *u* says that P, then *u* is true if and only if P.

(F)   If *u* says that P, then *u* is false if and only if not P.

The rationale for (T) and (F) is straightforward: if an utterance says, for example, that Jones is bald, then it is true if and only Jones is bald and false if and only if Jones is not bald (p. 190). Williamson formulates the principle of bivalence for utterances that "say something" as follows:

(B)   If *u* says that P, then either *u* is true or *u* is false (p. 187).

The restriction of bivalence to utterances that "say something" is intended to handle cases in which an utterance of a meaningful *sentence*, such as "This dagger is sharp", does not say anything, and so is neither true nor false, because the singular term "This dagger", as used on that occasion, has no reference.

Williamson's argument then proceeds as follows. Suppose (toward a contradiction) that *u* is a counterexample to (B). Then for some P, *u* says that P, *u* is not true, and *u* is not false. Since *u* says that P, from (T) and (F) it follows that *u* is true if and only if P, and *u* is false if and only if not P. Since *u* is not true and *u* is not false, we can conclude that not P and not not P, a contradiction. Williamson concludes that "it is coherent to suppose [vague utterances] to be neither true nor false only at the cost of treating them as though they said nothing" (p. 198).

The apparent force of this argument depends on an equivocation between a naive and a sophisticated understanding of phrases of the form '*u* says that P', where 'P' replaces a sentence. This becomes clear when we consider the liar paradox. We would naturally say that an utterance *u* of 'This utterance is not true' says that *u* is not true. If we let 'P' stand for '*u* is not true', then we are inclined to say that *u* says that P. But by (T) it follows that *u* is true if and only if *u* is not true. Williamson assumes that (T) elucidates the notion of truth in (B) (p. 188). Hence the consequence that *u* is true if and only if *u* is not true would be a clear counterexample to (B), if that principle were meant to apply to our naive sense of '*u* says that P'. But Williamson simply *denies* that *u* is a counterexample to (B),

and so he concludes that "*u does not say* that *u* is not true" (p. 197). This shows that to accept (B), we must adopt a sophisticated interpretation of the phrase '*u* says that P'.[10]

Williamson argues that if a *vague* utterance "says something" in a naive sense, it must also "say something" in the sophisticated sense, because it is composed of words we understand in combinations we understand (p. 196). But the liar paradox is also expressed by using words we understand in combinations we understand well enough to see that it is contradictory. Hence the apparent compositional structure of a sentence does not justify Williamson's conclusion that if a vague utterance "says something" in a naive sense, it "says something" in the sophisticated sense. Williamson also claims that for any borderline vague utterance, "we can envisage circumstances in which the utterance would have been clearly true while saying just what it actually says" (p. 196). But this claim just *presupposes* that such a borderline vague utterance actually "says something" in the sophisticated sense. Hence the claim does not offer any independent support for Williamson's conclusion that if a vague utterance "says something" in a naive sense, it "says something" in the sophisticated sense.

I conclude that we can accept (B) and the sophisticated interpretation of the phrase '*u* says that P' that goes with it *without* giving up the intuition that there are some vague utterances that "say something" in a naive sense of that phrase, but are neither true nor false, and so Williamson's argument really does not touch the intuition that many vague utterances that "say something" are neither true nor false. Nevertheless, to highlight my doubts about Williamson's explanation of our ignorance of borderline vague truths, I will assume with Williamson that all vague utterances that "say something" are either true or false.

5. WILLIAMSON ON INEXACT KNOWLEDGE, RELIABILITY, AND MARGINS FOR ERROR

According to Williamson's epistemic theory of vagueness, there are many vague utterances whose truth values we cannot know. Perhaps the biggest obstacle to accepting this is the feeling that we should be able to discover the truth value of a vague utterance if it has one. Williamson tries to remove this obstacle by developing a theory that explains why we cannot know the hypothetical sharp boundaries of our vague concepts.

His theory depends on an analogy between our knowledge of vague truths and what he calls *inexact knowledge*: the kind of knowledge I can have of the number of people in a stadium, for example, when I *see* a crowd of people in the stadium, but I do not take the time to count them.

Williamson assumes that "a belief constitutes knowledge only if it is reliable enough" (p. 226). If my eyesight and ability to judge numbers are normal, and the stadium is of average size for football, then my belief that there are not exactly two hundred or two hundred thousand people in the stadium is reliable enough to constitute knowledge. On the other hand, for many numbers n, I do not know that there are not exactly n people in the stadium.

Williamson theorizes that a belief that there are not exactly n people in the stadium is reliable enough to constitute knowledge only if in all sufficiently similar cases, there are not exactly n people in the stadium, hence only if n leaves a *margin for error* sufficient for knowledge of how many people there are in the stadium. If k is the actual number of people in the stadium, then n leaves margin for error m if and only if $|n - k| \geq m$. If m is sufficient for knowledge of how many people there are in the stadium, surely $m > 1$, so we can accept

(10)   If I know that there are not exactly n people in the stadium, then there are not exactly $n + 1$ or $n - 1$ people in the stadium.

Williamson uses the idea of margins for error to distinguish between true belief and knowledge. Suppose I believe that there are not exactly n people in the stadium, and this belief is true, but n does not leave a margin for error sufficient for knowledge. Then there is a similar situation in which there are exactly n people in the stadium but I nevertheless believe that there are not exactly n people in the stadium, and so my actual true belief is not reliable enough to constitute knowledge. On the other hand, if n does leave a margin for error sufficient for knowledge, then in all sufficiently similar cases in which I judge that there are not exactly n people in the stadium, there are not exactly n people in the stadium, and so my actual belief is reliable enough to constitute knowledge.[11]

6. WILLIAMSON'S EXPLANATION OF OUR KNOWLEDGE AND
IGNORANCE OF SHARP BOUNDARIES AND BORDERLINE TRUTHS

Assuming that there are substantive supervenience conditions for vague concepts, Williamson develops a theory that explains why we cannot know the hypothetical sharp boundaries of our vague concepts. The basic idea is our knowledge of sufficient conditions for applying vague concepts depends on margin for error principles similar to the margin for error principles on which our inexact knowledge depends. Suppose k is the least number such that any person with k hairs on his head is bald, and any

person with $k + 1$ hairs on his head is not bald.[12] Then my belief that a person with n hairs on his head is bald leaves margin for error m if and only if $(k - n) \geq m$. If m is sufficient for knowledge that a person with n hairs on his head is bald, surely $m > 1$, so we can accept

(11)   If I know that a person with n hairs on his head is bald, then a person with $n + 1$ hairs on his head is bald.

Unlike (10), however, (11) is motivated by an analysis of reliability that is special to judgments about sufficient conditions for applying vague concepts. The key idea behind this account of reliability is that the meanings of vague terms supervenene on the use of those terms in our linguistic community in such a way that a small change in our dispositions to use a given vague term would amount to a small change in the meaning and extension of that term. According to Williamson, my actual assertion of the sentence 'A person with n hairs on his head is bald' is the outcome of a disposition to be reliably right only if my assertion of that sentence would have expressed a truth even if our overall use of 'bald' had been slightly different.[13]

Williamson's idea that some changes in the use of a vague term are "small" and others are not is meant to clarify the initially attractive idea that the *source* of our ignorance of the truth values of borderline vague claims is "conceptual" not "empirical".[14] His account of reliability for vague judgments is only as clear as his idea that some changes in the use of a vague term are "small" and others are not. For reasons I will explain below, this idea is the Achilles heel of his proposed explanation of our ignorance of borderline vague truths.

If we know the relevant margin for error principles, we can use Williamson's theory to explain why we must remain ignorant of the hypothetical sharp boundaries of our vague concepts. Suppose, as before, that k is the least number such that any person with k hairs on his head is bald, and any person with $k + 1$ hairs on his head is not bald; then k marks a sharp boundary for bald. But my belief that any person with k hairs on his head is bald does not leave a margin for error m sufficient for knowledge, since $(k - k) = 0$, and m is sufficient for knowledge only if $m > 0$. Thus we cannot know such hypothetical boundaries because they do not leave a margin for error sufficient for knowledge. In the case of inexact knowledge, there is no theoretical barrier to our obtaining exact knowledge; in principle, we could always count the number of people in the stadium, for example. But when it comes to our knowledge of sufficient conditions for applying vague concepts, according to Williamson, we cannot overcome the limits imposed by margin for error principles such as (11).

If we know that (9) is true and that (11) is the margin for error principle relevant to evaluating knowledge-claims made by using the term 'bald', then we can use Williamson's theory to explain why we do not know that x is bald, where x is a borderline case of bald. In principle, by counting we can learn that the number of hairs on x's head is n. If we know that x is bald and that the number of hairs on x's head is n, then from (9) we can infer that every person with n hairs on his head is bald. Together with (11), this implies that every person with $n+1$ hairs on his head is bald. But since x is a borderline case of bald, it may be that not every person with $n+1$ hairs on his head is bald; our disposition to assert that x is bald is not reliable enough for knowledge, and so we do not know that x is bald.

7. TWO ROLES FOR THE ASSUMPTION THAT THERE ARE SUBSTANTIVE SUPERVENIENCE CONDITIONS FOR VAGUE CONCEPTS

In this explanation there are two roles for the assumption that there are substantive supervenience conditions for vague concepts. The first and most important role of supervenience principles like (9) is to express conceptual or metaphysical relationships between properties that (we take to) settle the truth or falsity of margin for error principles like (11). Given (9), one can support (11) as follows. The actual assertion of a true sentence of the form 'everyone with n hairs on his head is bald' is the outcome of a disposition to be reliably right only if that sentence would still have expressed a truth in circumstances that involve only a "small" difference in our overall use of 'bald'. It seems plausible to start by assuming that a difference in our overall use of 'bald' is "small" only if it does not result in sweeping changes in the truth values of sentences about whose truth values we are very confident. This constraint clearly shows that our understanding of what counts as a "small" change in the use of a vague term is not independent of our confidence in knowledge claims we express by using that term. Now suppose in addition that a difference in our overall use of 'bald' is "small" only if despite that difference in use, (9) still expresses a truth.[15] Given these constraints, a change in the extension of 'bald' is small only if it is confined to *borderline* cases of 'bald', for any other changes in the extension of 'bald' would require sweeping changes in the truth values of sentences about whose truth values we are very confident. It is then plausible to conclude that the actual assertion of 'everyone with n hairs on his head is bald' expresses knowledge only if n is at least 1 hair away from the hypothetical sharp boundary for 'bald',[16] hence only if 'everyone with

$n + 1$ hairs on his head is bald' expresses a truth. This is just what (11) says.

But if we do not know any conceptual or metaphysical relationship, such as (9), between being bald and having some other property, then for all we know *every* true sentence of the form 'x is bald if and only if y is bald' might easily have been false. To see this, we need only suppose that a change in the use of 'bald' is "small" if the resulting extension of 'bald' differs from its actual extension by just one individual.[17] By this criterion, Williamson's theory of reliability for vague judgments would imply that for *every n*, if 'everyone with n hairs on his head is bald' expresses a truth, someone might very easily have made a false assertion with that sentence, since a "small" change in our use of 'bald' might have made the extension of 'bald' differ from its actual extension by just one individual with n hairs on his head. This would undermine our supposed *knowledge* that for some n, everyone with n hairs on his head is bald, so it would make (11) at best only vacuously true. I will return to this argument below.

The second role of the assumption that there are substantive supervenience conditions for vague concepts is to link beliefs expressed by sentences of the form 'x is F', where 'F' is a vague term, to margin for error principles such as (11). To explain why we must be ignorant of whether x is bald if x is a borderline case of baldness, we suppose that if we know that x is bald, then we know that everyone with the same number of hairs on his head is also bald; then we use Williamson's explanation of why we cannot know that everyone with that many hairs on his head is bald to infer that we do not know that x is bald.

8. DOUBTS ABOUT THE EXISTENCE OF SUBSTANTIVE SUPERVENIENCE CONDITIONS FOR VAGUE CONCEPTS

But why should we accept the assumption that there are substantive supervenience conditions for vague concepts? When we gave up the idea that supervenience conditions are fixed by meaning postulates or constraints on valuations, we deprived ourselves of the most promising way of defending that assumption. We saw earlier that the existence of substantive supervenience conditions, which amounts to the existence of sharp boundaries in the strong sense for vague concepts, does not follow just from the acceptance of bivalence for vague utterances.

Williamson himself stresses that even if meaning supervenes on use, the relation between meaning and use is difficult to survey. At one point he observes that "meaning may supervene on use in an unsurveyably chaotic

way" (p. 209). This should lead us to doubt that there are substantive supervenience conditions for vague concepts.

To illustrate this point, let us briefly reconsider (1a), the statement that a person with 0 hairs on his head is bald. Consider Gandhi. I believe that he lost all his hair as a result of a natural process. I would say that in the last years of his life he was bald. It is tempting to generalize from this case to the conclusion that *every* person with 0 hairs on his head is bald. But what about Michael Jordan? There is no hair on his head, but is he bald? Suppose he shaves his head every day, and if he did not shave his head, his hair would grow back. Then I would say that Michael Jordan shaves his head, but he is not bald.

You will be understandably irritated by this example if you take for granted that we can *stipulate* that if a person has no hair on his head, then he is bald, whether or not he shaved his head. But if we reject the use of meaning postulates or constraints on valuations, as Williamson rightly does, then we cannot simply stipulate the meanings of vague terms; we must investigate their use and arrive at a reasonable account of what they mean.

Another possible reaction is that we have not arrived at a proper specification of sufficient conditions for being bald. Perhaps we should say that any person who lost all his hair as a result of a natural process is bald. But even if we find this characterization more plausible, the Jordan example suggests that no statement of sufficient conditions for applying a vague concept is analytic or immune to counterexamples. If there are true generalizations of this form, they are deeply entrenched generalizations, not logical truths. And (9) may be false even if 'anyone with n hairs on his head is bald' is true for many values of 'n'. For any n, whether or not a person with n hairs on his head is bald may be related in "an unsurveyably chaotic way" to an open-ended range of factors, including the size and shape of his head, the way his hair is distributed on his head, the thickness of his hair, its color and type (straight, wavy, curly), his age, and so on.

This casts doubt on the assumption that there are substantive supervenience conditions for the concept bald. Similar considerations cast doubt on the assumption that there are substantive supervenience conditions for other vague concepts, including those expressed by such terms as 'heap', 'thin', and 'old', to name just a few. But if we do not know of any substantive supervenience conditions for these vague concepts, we cannot use Williamson's theory to explain why we should be ignorant of borderline truths expressed by sentences of the form 'x is F', where 'F' is one of these vague terms.[18]

9. FOUR REPLIES

One might grant that Williamson's theory explains our ignorance in borderline cases only if there are substantive supervenience conditions for vague concepts, but deny that to use Williamson's theory to explain our ignorance we must *know* that there are substantive supervenience conditions for vague concepts. One might claim that to use Williamson's theory to explain our ignorance in borderline cases, we need only *believe* that there are substantive supervenience conditions for vague concepts. If this claim is correct, then to discredit Williamson's theory, I must prove that there are no substantive supervenience conditions for vague concepts. But I have only raised *doubts* about whether there are substantive supervenience conditions for vague concepts; I have not proved that there are no substantive supervenience conditions for vague concepts, and the kinds of doubts that I did raise suggest that I cannot prove this.

This reply rests on the crucial claim that to use Williamson's theory to explain our ignorance in borderline cases, we need only *believe* that there are substantive supervenience conditions for vague concepts. But the reply begins by *granting* my conclusion that Williamson's theory explains our ignorance in borderline cases only if there are substantive supervenience conditions for vague concepts. This conclusion implies that a Williamson-style "explanation" of our ignorance in borderline cases is not a *true* explanation *unless* there are substantive supervenience conditions for vague concepts. The conclusion therefore *implies* that to use Williamson's theory to explain our ignorance in a borderline case of a particular vague concept, we must *know*, or at least *have very good reason to believe*, substantive supervenience conditions for applying that concept. I observed that Williamson's own view of the relationship between meaning and use should undermine any confidence we might have had in the claim that there are substantive supervenience conditions for vague concepts. With our confidence in this claim undermined, there is no basis for using Williamson's theory to explain our ignorance borderline cases.

One might reply that Williamson's explanation of our ignorance of borderline truths does not require that there be substantive supervenience conditions for vague concepts, such as the condition specified by (9). Suppose that Jones has been losing his hair by a natural process, and that we now would say he is bald. One might argue that

(12)   If we know that Jones is bald when he has n hairs, then he was bald when he had $n + 1$ hairs.

because otherwise our judgement 'He is bald' when he has n hairs would not be based on a sufficiently reliable disposition. If Jones is a borderline

case of bald when he has n hairs, it may very well be that he was *not* bald when he had $n + 1$ hairs. Hence if we accept (12), we must conclude that the judgment that Jones is bald when he has n hairs is not reliable enough for knowledge. In this explanation of ignorance, there is no appeal to (9) or to any other substantive supervenience condition on (or for) being bald.[19]

I grant that if we had reason to accept (12), this *would* be an acceptable explanation of our ignorance of whether Jones is bald when he has n hairs. But what reason do we have to accept (12)? Suppose that we carefully *count* the number of hairs on Jones's head, and we claim to know that Jones is bald when he has n hairs, but not when he had $n + 1$ hairs. Why isn't this just like the case of someone who claims to know after a careful count that there are 23,890 people in the stadium, but that there are not 23,891 people in the stadium? The answer cannot be that knowledge requires reliability.

One might reply that when we assert 'Jones is bald when he has n hairs, but not when he had $n + 1$ hairs', even if we speak truly, the use of "bald" could easily have been such that we would have spoken falsely, and so we are too unreliable to know.[20] The trouble is that if we don't know of any conceptual or metaphysical relationships between being bald and having some other property, we have no basis for saying whether a change in the use of 'bald' is "small", and so we have no basis for accepting this reply. To see this, suppose again that a change in the use of 'bald' is "small" if the resulting extension of 'bald' differs from its actual extension by just one individual. By this criterion, Williamson's theory would imply that for every n, we do not know that Jones is bald when he has n hairs on his head. This would undermine our supposed *knowledge* that for some n, Jones is bald when he has n hairs on his head. But what we wanted was an explanation of why we do not know *borderline* vague truths, not a theory that implies that we do not know *any* vague truths!

To this objection one might reply that we know *intuitively* whether or not a change in the use of 'bald' is "small", even though we cannot say *how* we know this. But in the context of Williamson's theory, an appeal to our intuitions about whether or not a change in the use of 'bald' is "small" is doubly problematic. First, unlike our intuitions about the ordinary counter-factuals we rely on in practical reasoning – such counterfactual as "Clinton might have lost the 1992 Presidential election" – we have no established practice of evaluating such counterfactuals as "the actual assertion of a true sentence of the form 'Jones is bald' would still have expressed a truth in circumstances that involve only a 'small' difference in our overall use of 'bald' " Second, any such intuitions, to the extent that we have them at all, merely reflect our intuitions about whether we know that x is bald. The principle behind such intuitions might be stated as follows: a difference in

our overall use of 'bald' is *small* only if it does not result in *any* changes in the truth values of utterances whose truth values we are confident that we know. If we accept this principle, then Williamson's explanation of our ignorance of whether *x* is bald ultimately just reflects our degree of confidence in the knowledge claim that *x* is bald, and so it is no more than an elaborate *restatement* of what it is supposed to explain.

10. CONCLUSION AND METHODOLOGICAL MORAL

I conclude that Williamson's proposed explanation of our ignorance of borderline vague truths faces a dilemma: either we have no grounds for accepting it, or it is no more than an elaborate restatement of what it is supposed to explain. The initial impression that we have grounds for accepting it is created by the false assumption that we know substantive supervenience principles for vague concepts. Supervaluationists use meaning postulates and constraints on valuations to build supervenience principles for concepts expressed by vague terms into the "logic" of our language. Williamson's view of the relation between meaning and use leads him to reject this approach. As I explained above (in Section 3), I agree with Williamson to this extent: in my view, a realistic description of the relation between meaning and use undermines the positivists' analytic-synthetic distinction and the supervaluationist's corresponding assumption that natural languages have built in semantical rules and constraints on admissible valuations. But I have argued that a realistic description of the relation between meaning and use should *also* lead us to reject the supervenience principles that Williamson assumes in his paradigm explanations of our ignorance of borderline vague truths. Although his theory does not officially rest on such supervenience principles, without them either we have no grounds for deciding whether a change in the use of a vague term is "small", or we just stipulate that a difference in our overall use of a vague term is "small" only if it does not result in *any* changes in the truth values of sentences whose truth values we are very confident that we know. But if we have no grounds for deciding whether a change in the use of a vague term is "small", then we have no grounds for accepting Williamson's proposed explanation of our ignorance of borderline vague truths. And if we just stipulate that a difference in our overall use of a vague term is "small" only if it does not result in any changes in the truth values of utterances whose truth values we are confident that we know, then Williamson's proposed explanation of our ignorance of borderline vague truths is at best an elaborate restatement of our starting observation that

we are not confident that we know the truth values of borderline vague utterances.

Quine once remarked that "meaning is what essence becomes when it is divorced from the object of reference and wedded to the word".[21] In a similar spirit, I would say that Williamson's intuitions about reliability are what meaning postulates and constraints on valuations for vague terms become when they are divorced from semantics and wedded to the world. As we have seen, the difficulty for Williamson's approach is that we have no grip on the supposedly "conceptual" sources of our ignorance, such as whether a particular change in the use of 'bald' is "small", apart from our "empirical" knowledge of who is bald. The methodological moral is that once we give up the supervaluation approach to vagueness, we must also give up the initially attractive idea that we can distinguish between "conceptual" and "empirical" sources of our ignorance of borderline vague truths.[22]

NOTES

[1] Timothy Williamson, *Vagueness* (London: Routledge, 1994). All page references in the text of this paper will be to this book.

[2] We can get a version of the original sorites argument for 'heap' by replacing 'P( )' with 'a collection of ( ) grains of sand is not a heap', as follows:

(1b)   A collection of 0 grains of sand is not a heap,

(2b) $\forall n$[(a collection of n grains of sand is not a heap) $\rightarrow$ (a collection of $n + 1$ grains of sand is not a heap)],

(3b)   A collection of 100,000 grains of sand is not a heap.

Like all sorites arguments, this one is paradoxical because the conclusion seems clearly false, yet both premises seem true. We feel confident that a collection of 0 grains of sand is not a heap, and we seem forced to accept the second premise, since we feel sure that 1 grain of sand cannot make the difference between being a heap and not being a heap.

[3] It is easy to show that (7) and (9) together imply (8), by deriving a contradiction from the conjunction (of regimented versions) of (7), (9), and the negation of (8), as follows (using the "Main Method" from *Quine's Methods of Logic*):

(1) $\forall x \forall y \forall n$[(($x$ is a person) $\wedge$ (x has n hairs on x's head) $\wedge$ (y is a person) $\wedge$ (y has n hairs on y's head)) $\rightarrow$ ((x is bald) $\leftrightarrow$ (y is bald))]. [paraphrase of (9)]

(2) $\exists x$[(x is a person) $\wedge$ (x has $n+1$ hairs on x's head) $\wedge$ $\neg$(x is bald) [paraphrase of (7)]

(3) $\exists x$[(x is a person) $\wedge$ (x has $n + 1$ hairs on x's head) $\wedge$ (x is bald)]   [neg. of paraphrase of (8)]

(4) $(z$ is a person$) \wedge (z$ has $n + 1$ hairs on z's head$) \wedge \neg(z$ is bald$)$ [EI from (2)]

(5) $(w$ is a person$) \wedge (w$ has $n + 1$ hairs on w's head$) \wedge (w$ is bald$)$ [EI from (3)]

(6) $((z$ is a person$) \wedge (z$ has $n + 1$ hairs on z's head$) \wedge (w$ is a person$) \wedge (w$ has $n + 1$ hairs on w's head$)) \rightarrow ((z$ is bald$) \leftrightarrow (w$ is bald$))]$.   [UI x3 from (1)]

Lines (4)–(6) are together truth-functionally inconsistent, so (1)–(3) are together inconsistent.

4 See Rudolf Carnap: 1952, 'Meaning Postulates', *Philosophical Studies* **3**, 65–73, reprinted in Rudolf Carnap, *Meaning and Necessity*, second edition, University of Chicago Press, Chicago, pp. 222–229.

5 See Kit Fine: 1975, 'Vagueness, Truth and Logic', *Synthese* **30**, 270.

6 For one version of this approach, see Carl Hempel: 'Vagueness and Logic', *Philosophy of Science* **6**, 163–180. In 'The Sorites Paradox and Higher Order Vagueness', *Synthese* **85**, 417–474, J. A. Burgess presents an account of higher order vagueness that bears some similarities to the one I recommend, although Burgess does not explicitly draw a distinction between precise languages and empirical semantics, and this weakens his theory. David Lewis draws the distinction in his paper 'Languages and Language', reprinted in his *Philosophical Papers*, Volume I, pp. 163–188. The distinction between precise languages and empirical semantics can be traced back to Carnap's distinction between pure and descriptive semantics.

7 See David Lewis, 'Languages and Language', p. 188.

8 See W. V. Quine: 1961, 'Two Dogmas of Empiricism', reprinted in W. V. Quine, *From a Logical Point of View*, revised second edition, Harvard University Press, Cambridge, MA, pp. 20–46; 'Carnap and Logical Truth', in Paul A. Schilpp (ed.), *The Philosophy of Rudolf Carnap*, La Salle, Illinois: Open Court, 1963, pp. 385–406; Chapter 2 of Word and Object MIT Press, Cambridge, MA, 1960.

9 See Hilary Putnam: 1975, 'The Analytic and the Synthetic', reprinted in Hilary Putnam, *Mind, Language, and Reality*, Cambridge University Press, Cambridge, pp. 33–69.

10 Note also that even if all vague utterances that "say something" in Williamson's sophisticated sense are either true or false, there remain other grammatical sentences of English, such as "This dagger is sharp", an utterance of which may be neither true nor false. So even if there are not any counterexamples to (B), some classical rules of inference – such as the rule that allows us to infer '$\exists y((y$ is a dagger$) \wedge (y$ is sharp$))$' from '$(z$ is a dagger$) \wedge (z$ is sharp$)$', a regimented version of "This dagger is sharp" – do not apply to all utterances of "This dagger is sharp" that seem to "say something" in Williamson's sophisticated sense. A complete vindication of (B) would not tell us when to use classical rules of inference for English.

11 Williamson also uses the idea of margins for error to explain why I may fail to know that I know that there are not exactly n people in the stadium. He claims that inexact knowledge obeys margin for error principles of the form " 'A' is true in all cases similar to cases in which 'It is known that A' is true", and that a special case of inexact knowledge is my knowledge of my inexact knowledge. Hence there is a margin for error relevant to evaluating the reliability of my belief that I know that there are not exactly n people in the stadium. Williamson notes that even if the margin for error relevant to evaluating this belief is the same as the margin for error relevant to evaluating my belief that there are not exactly n people in the stadium, it may be that I know that there are not exactly n people in the stadium, but I do not know that I know this.

[12] We take (9) for granted, so if someone with $k + 1$ hairs on his head is not bald then everyone with $k + 1$ hairs on his head is not bald.

[13] Williamson's account of linguistic competence is anti-individualistic, in the sense that members of a linguistic community need not have exactly the same dispositions to be counted as expressing the same vague concepts when they use the same vague terms of their shared language. Vagueness is a source of inexactness, Williamson claims, because "individual uses of a vague term are not fully sensitive to small differences in the overall pattern on which small differences in meaning supervene" (p. 235). Moreover, Williamson's account of reliability for vague utterances is compatible with our knowledge of the actual meanings expressed by our vague terms, since to know the meaning of a vague term, "it is enough to use the term within the appropriate practice" (p. 237). (11) does not lead to a sorites paradox for 'known to be bald', since, as we saw in note 12, if my knowledge is subject to margins for error, I may know something yet fail to know that I know it.

[14] "Conceptual Sources of Exactness" is the revealing title of Section 8.4 (*Vagueness*, pp. 230–234), in which Williamson presents his explanation of our ignorance of borderline vague truths.

[15] This constraint is analogous to the supervaluationist's assumption that (9) is a constraint on all admissible specifications of the extension of 'bald'.

[16] To simplify matters, I assume that there is exactly one sharp boundary (in the strong sense) for 'bald'.

[17] Note that no such minimal change in extension would conflict with the principle from the previous paragraph that a difference in our overall use of 'bald' is small only if does not result in sweeping changes in the truth values of sentences about whose truth values we are very confident.

[18] In "Vagueness, Ignorance, and Margin for Error" (*British Journal for the Philosophy of Science* **46**, 1995, 589–601), R. M. Sainsbury's review of Williamson's book *Vagueness*, Sainsbury notes that "the way in which Williamson describes the relevant margin for error mechanisms presupposes ... supervenience [according to which vague facts supervene on precise ones]: worlds differing in a vague respect must differ in a precise one" (p. 593, footnote 2, Sainsbury credits Michael Martin for the observation). Sainsbury does not pursue the question of whether doubts about supervenience claims would undermine Williamson's use of margin for error principles to explain our ignorance of borderline vague truths.

[19] This reply is due to Williamson.

[20] This is Williamson's explanation and defense of the previous reply.

[21] W. V. Quine, 'Two Dogmas of Empiricism', in *From a Logical Point of View*, p. 22.

[22] Thanks to Scott Kimbrough, Tom Meyer, Miriam Solomon, Joan Weiner, and especially Timothy Williamson for helpful comments on earlier drafts, and to Scott Weinstein for an illuminating discussion of the sorites paradox.