An Anti-Epistemicist Consequence of
Margin for Error Semantics for Knowledge

DELIA GRAFF
Princeton University
Forthcoming in *Philosophy and Phenomenological Research*

1 Transparent Propositions and Margin for Error Semantics

Let us say that the proposition that p is transparent just in case it is known that p, and it is known that it is known that p, and it is known that it is known that it is known that p, and so on, for any number of iterations of the knowledge operator 'it is known that'. If there are transparent propositions at all, then the claim that any man with zero hairs is bald seems like a good candidate. We *know* that any man with zero hairs is bald. And it also does not seem completely implausible that we know that we know it, and that we know that we know that we know it, and so on.

Mario Gómez-Torrente (1997, p. 244) observes that if $B(0)$ is transparent (where in general, $B(n)$ stands for the claim that any man with n hairs is bald), then not all of the following of Timothy Williamson's margin for error principles can be true.

Williamson's margin for error principles:

$(1^{(m)})$ $(\forall n)(K^{m+1}B(n) \longrightarrow K^m B(n+1))$

- K^m abbreviates m iterations of 'it is known that'.

- $(1^{(m)})$ is a margin for error principle for $K^m B(x)$.

For if the margin for error principles were all true, then given the transparency of $B(0)$—in particular, the truth of $K^{10^6}B(0)$—by one-million of the margin for error principles (namely $(1^{(0)})$–$(1^{(10^6-1)})$) it would follow that any man with one-million hairs is bald, which is false. The question arises whether rejection of the margin for error principles is an option for Williamson, since he argues that margin for

error principles are what explain the ignorance postulated by epistemicist theories of vagueness.[1]

In response to Gómez-Torrente, Williamson (1997, 262–63) concedes for the sake of argument that $B(0)$ may be a transparent proposition. He proceeds to provide a possible worlds model of our knowledge about baldness, in which $B(0)$ *is* a transparent proposition, but in which his explanation of our ignorance of the cut-off point for 'bald' can in essence still be retained.

Williamson's (1997, 262–63) Model:

  (i) $W = \{s_n : n \in \mathbb{N} \ \& \ 1 \leq n\}$

  (ii) $s_i \models B(n)$ iff $n < i$

  (iii) $s_i R s_j$ iff $|i - j| \leq 1$

  (iv) $s_i \models KA$ iff $(\forall s_j)(s_i R s_j \longrightarrow s_i \models A)$

The structure of Williamson's model is represented by the following diagram:

$$s_1 \longleftrightarrow s_2 \longleftrightarrow s_3 \longleftrightarrow s_4 \cdots$$
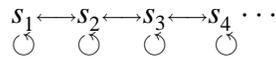
Here we have a series of worlds. Clause (ii) tells us that at the first world in the series the cut-off for B occurs between 0 and 1, and then shifts one hair to the right at each successive world in the series, reflecting the slight differences in meaning the predicate B might have. The third and fourth clauses here reflect Williamson's idea that knowledge requires a margin for error: that if the judgement expressed by a certain sentence is to constitute knowledge, then that sentence must be true at worlds very similar to our own, in particular at worlds in which the sentence has just a slightly different meaning; for if the sentence is not true at worlds very similar to our own, then our judgement cannot have been *reliable*, since the judgement would have been mistaken had things been slightly different.

$B(0)$ is transparent at each world in this model since by (i) and (ii), $B(0)$ is true at every world in the model, and so by (iii) and (iv), $K^m B(0)$ is true at every world in the model, for any m. Accordingly, by our previous result, there is no world in the model at which all of

---

[1]See *Vagueness* (Williamson 1994) ch. 8.

Williamson's margin for error principles can be true. But still, it is the case that at each world in the model B has an unknown sharp cut-off point. For each $i > 0$:

$$s_i \models B(i-1) \;\&\; \neg B(i) \;\&\; \neg K(B(i-1) \;\&\; \neg B(i)).$$

In fact, we also have that the following restricted versions of Williamson's margin for error principles are still all true at each world in the model:

$(1'^{(m)})$ $(\forall n > 0)(K^{m+1}B(n) \longrightarrow K^m B(n+1))$

In conceding that there may be transparent propositions, Williamson is in effect conceding that his margin for error principles might require some restriction, but he maintains that they nevertheless provide the required support for his epistemicism.

This last claim of Williamson's comes under attack from a variety of directions in Gómez-Torrente's "Vagueness and Margin for Error Principles." Here I focus on one line of that attack, Gómez-Torrente's argument that Williamson's margin for error semantics has a formal consequence that is inconsistent with his epistemicism about vagueness.

In §3.2 of "Vagueness and Margin for Error Principles" Gómez-Torrente shows that at each world in Williamson's (1997, 262–63) model, we have the following true for some sufficiently large m:[2]

$$K(K^m B(0) \;\&\; \neg K^m B(1)).$$

Since one of the worlds in the model is supposed to represent the actual world (though we do not know which world that is), according to the model there is a predicate $K^m B(x)$ whose cut-off point is actually known. If this predicate is a *vague* one, then Williamson must have been wrong to concede the transparency of $B(0)$, since then given Gómez-Torrente's result, we would have a vague predicate whose cut-off point is known, that is, a vague predicate for which Williamson's theory is false!

But we should not be so quick. Williamson had presented us with a simple possible worlds model in order to demonstrate that he could accommodate the transparency of $B(0)$ while retaining in essence his

---

account of our ignorance of the cut-off point for 'bald'. His simple model, however, is surely not accurate. Certain assumptions were made for the sake of concreteness: in particular, it was assumed that the margin for error required for knowledge about baldness is a margin of just one hair. Other certainly false assumptions were made for the sake of simplicity: in particular, it was assumed that there is just one dimension of variation relevant for the applicability of 'bald'— variation in number of hairs. The fact that in this simple model there is an apparently vague predicate with a known cut-off point does not in and of itself present a problem for Williamson. What we especially want to know is whether Gómez-Torrente's result depends crucially on the simplifying assumptions, or whether it can be further generalized. As it turns out, more general versions of the result do hold. In the next section, I discuss the extent of this generalizability.

2   Generalizations of Gómez-Torrente's Result

In the appendix to his (1994) book, Williamson actually provides two alternative semantics for an operator 'it is clearly the case that'. I will here assume those semantic proposals to apply equally well to the knowledge operator 'it is known that'.[3] In each case, knowledge requires a certain margin for error. In the first version of the semantics, the margin for error is fixed.

2.1   Fixed-Margin Semantics for Knowledge

- A fixed-margin model is a quadruple $\langle W, d, r, \text{VAL} \rangle$.

- W is the (non-empty) set of possible worlds in the model.

- $d(w, v)$ is a non-negative real number that measures the similarity between any two worlds w and v in W.

  The measure may be thought of as a distance, since we may think of the worlds in a model as arranged in a space in such a way

---

[3]This is not to say that the two operators have the same meaning. Rather, it is just to say that at a certain level of generality, their meanings have the same logical structure. We might take it that the difference between the two, on a margin for error approach like that of Williamson's, concerns the magnitude of the margin for error required—that clarity requires a greater margin for error than mere knowledge. To model a language containing both operators, we might require each model to contain two different margins for error, one for each operator.

that the distance between two worlds in the space is a reflection of their similarity. The lesser the distance between worlds, the more similar they are. Accordingly, we stipulate that: the measure is symetric, that is, $d(w, v) = d(v, w)$; that $d(w, w) = 0$, since each world w is exactly similar to itself; and that the measure satisfies the triangle inequality, that is, $d(w, u) \leq d(w, v) + d(v, u)$.

- r is a positive real number representing the margin for error required for knowledge.[4]

- VAL is a valuation function that assigns the values true or false to pairs of sentences and worlds. We let $w \models A$ abbreviate VAL(A,w) is true. VAL meets the standard classical constraints, as well as the following clause for the knowledge operator K.

  VAL K: $w \models KA$ iff $(\forall v \in W)(d(w, v) \leq r \longrightarrow v \models A)$

  That is, KA is true at a world w, just in case A is true at every world whose similarity to w is within the margin for error r.

Now, in seeking to generalize Gómez-Torrente's result, we will question whether fixed-margin models have the following property:

Gómez-Torrente's Property (First Version):

If, at any world w, there are transparent and non-transparent propositions of the form $B(n)$, then for some n and some sufficiently large m we have:

$$w \models K(K^m B(n) \ \& \ \neg K^m B(n + 1)).$$

That is, $K^m B(x)$ is a predicate whose cut-off point is known at w to occur between n and $n + 1$.

If a model has this property, we will call it a "GT-model." Gómez-Torrente proves of one particular fixed-margin model that it is a GT-model—namely, the model proposed by Williamson, discussed above

---

[4]Williamson (1994, p. 270) requires only that the margin for error be non-negative. I will rule out models with a zero margin for error since here we are concerned only with models of a vague language.

in §1, with $W = \{s_n : n \in \mathbb{N}\ \&\ 1 \le n\}$, $d(s_i, s_j) = |i - j|$, and $r = 1$.

But as Gómez-Torrente also points out, given the fixed-margin semantics, it is not only the model proposed by Wiliamson that is a GT-model, but also each of a certain class of models referred to as "the cropped models." Let me now give a characterization of this class. An "uncropped" model contains a set of worlds $\{s_i : i \in \mathbb{N}\}$, a similarity measure $d(s_i, s_j)$ defined as the distance between i and j, a margin for error not less than 1, and a valuation function VAL such that $\text{VAL}(B(n), s_i)$ is true just in case $n < i$.

An Uncropped Model

- $W = \{s_i : i \in \mathbb{N}\}$

- $d(s_i, s_j) = |i - j|$

- $1 \le r$

- $s_i \models B(n)$ iff $n < i$

In an uncropped model, $B(0)$ is not transparent at any world. Since in an uncropped model there is a world (namely s_0) at which even the claim that any man with zero hairs is bald is false, we have that $s_0 \models \neg B(0)$, and so that $s_1 \models \neg KB(0)$, and so that $s_2 \models \neg KKB(0)$, and in general that $s_i \models \neg K^i B(0)$.

To obtain a model in which $B(0)$ is transparent at each world, we can "crop" s_0 from the set of worlds in an uncropped model, and restrict the valuation function accordingly. To obtain a model in which $B(0)$ and $B(1)$ are transparent at each world, we can crop s_0 and s_1 from an uncropped model, and restrict the valuation function accordingly. The more worlds we crop, the more transparencies we get. Call a model a *cropped model* just in case it is obtained from an uncropped model by removing some initial segment from the sequence of worlds in the uncropped model, and restricting the valuation function accordingly. (We allow only initial segments to be cropped in order to ensure that if $B(n)$ is transparent at a world in the model, then so is $B(l)$ for any $l < n$.)

Gómez-Torrente remarks that it can be shown of each cropped model that it is a GT-model via an inductive proof analogous to the one he provides for the special case already discussed. But it is also possible to show in one fell swoop that all of the cropped models are

GT-models, by showing that any arbitrarily chosen one is. Suppose that $\mathbf{M}_j^r$ is a cropped model with $s_j, s_{j+1}, s_{j+2}, \ldots$ as its sequence of worlds, and the margin for error r. (Since $\mathbf{M}_j^r$ is a cropped model, $1 \leq j$.) Then we can prove by induction on m that:

$$s_i \models_{\mathbf{M}_j^r} K^m B(n) \text{ iff } n < j \text{ or } rm + n < i \quad (\text{for any } s_i \in W).$$

Given this fact one can show that any cropped model $\mathbf{M}_j^r$ has Gómez-Torrente's property, since a bit of arithmetic plus VAL K yield the following:

$$s_i \models_{\mathbf{M}_j^r} K(K^m B(j-1) \,\&\, \neg K^m B(j)) \text{ iff } \tfrac{i-j}{r} + 1 \leq m$$
(for any $s_i \in W$).

2.2 Fixed-Margins Semantics and the Brouwer Schema

We may also show that any model whatsoever is a GT-model, given Williamson's fixed-margin semantics for K. As Williamson (1994, p. 271) points out, his fixed-margin semantics is equivalent to a standard accessiblity-relation possible worlds semantics, since we may treat K as the box of modal logic by letting

$$w \models KA \text{ iff } (\forall v)(wRv \longrightarrow v \models A),$$

and defining an accessibility relation R as follows:

$$wRv \text{ iff } d(w,v) \leq r.$$

The accessibility relation thus defined will be symmetric, since degree of similarity is by definition symmetric—that is $d(w,v) = d(v,w)$, for all $w, v \in W$. Given the symmetry of the accessibility relation, we have the following validity (the Brouwer schema):

$$\mathbf{B} \colon \models \neg A \longrightarrow K\neg KA.$$

But now Gómez-Torrente's property is but a short step away. For suppose that $B(n)$ is transparent at a world w, but that $B(n+1)$ is not transparent at w.[5]

---

[5] For generality, we can allow for the case in which objects in this sorites space are densely ordered by taking 'n' and '$n+1$' to be names for objects sufficiently close together so that a boundary between them would count as a sharp cut-off.

- That is, we have:

    (1) $w \models \neg K^m B(n+1)$, for some sufficiently large m, but

    (2) $w \models K^l B(n)$, for all l.

- Given the validity of the **B** schema, from (1) we get:

    (3) $w \models K \neg K^{m+1} B(n+1)$

- Given that $(KA \ \& \ KB)$ implies $K(A \ \& \ B)$, from (2) and (3) we get:

    (4) $w \models K(K^{m+1} B(n) \ \& \ \neg K^{m+1} B(n+1))$

In other words, given that $B(n)$ is transparent, but that $K^m B(n+1)$ is false at w, the cut-off point for $K^{m+1} B(x)$ is known at w to occur between n and $n+1$. Thus we have shown that every fixed-margin model, without qualification, is a GT-model. In particular, we have made no assumptions about the magnitude of the margin for error required for knowledge, or about the dimensions of variation relevant for the applicability of 'bald'.

There's a worry about Williamson's fixed-margin semantics for knowledge, however, namely, its validation of the **B** schema, according to which if a proposition is false, then it is known that it's *not* known to be true. But this doesn't seem right. For not only is it very likely that there is some false proposition p which we believe, it is also very likely that there is some false proposition p which we believe that we *know*. But were the analog of the Brouwer schema for knowledge to hold good, the falsity of the proposition that p would entail that we *know* that we don't know that p, and hence, presumably, that we *believe* that we don't know that p. Thus were the analog of the Brouwer schema for knowledge to hold good, then when we believe that we know some proposition p, the falsity of p suffices to render our beliefs about what we know inconsistent. Even worse (perhaps), if knowledge entails belief, then the analog of the Brouwer schema requires that for every false proposition, we have a true belief about whether we know it.

Williamson briefly expresses some worry about the **B** schema for the knowledge operator. In fact, he offers an alternative semantics for knowledge on which it is not validated.

2.3   Variable-Margin Semantics for Knowledge

- As with a fixed-margin model, a variable-margin model is a quadruple $\langle W, d, r, \text{VAL} \rangle$, where W is the (non-empty) set of possible worlds in the model, $d(w, v)$ is a non-negative real number that measures the similarity between any two worlds w and v in W, and r is a positive real number representing a margin for error.[6]

- The valuation function is as before except for the clause for K.

  VAL K: $w \models KA$ iff $(\exists \delta > 0)(\forall v \in W)(d(w, v) \leq r + \delta \longrightarrow v \models A)$
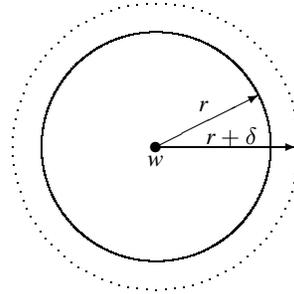
  That is, KA is true at a world w, just in case there's *some* margin for error $r + \delta$ greater than r, such that A is true at every world whose similarity to w is within the margin for error $r + \delta$.

Fixed-Margin Semantics   **Variable-Margin Semantics**
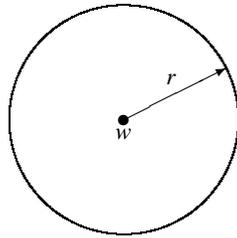
$w \models KA$ iff A is true at every world in the circle centered around w with radius r.

$w \models KA$ iff A is true at every world in *some* circle centered around w with a radius greater than r.

Now, it turns out that the **B** schema, which we appealed to in showing that every *fixed*-margin model is a GT-model, is not validated by the *variable*-margin semantics for knowledge. Of course, this fact

---

[6] Again (see also nt. 4), Williamson (1994, p. 272) allows models with zero-value margins for error, but we will rule them out here.

alone does not suffice to show us that not all variable-margin models are GT-models. As it turns out, however, they are not.

But still, as Gómez-Torrente notes,[7] every cropped model, as defined above, is a GT-model, even on Williamson's variable-margin semantics. This is so for an uninteresting reason, however. In the cropped models, where the worlds in the model can be arranged in a sequence, each a fixed distance (in terms of similarity) of 1 from the next, the fixed-margin and variable-margin semantics are simply equivalent.[8]

In fact, the fixed-margin and variable-margin semantics will come apart only in models in which the distance between worlds can be greater than the margin for error r without having to be greater than r by at least some given amount—that is, only if the set

$$\{x : (\exists w)(\exists v)(d(w, v) = x) \ \& \ x > r\}$$

has r itself as a greatest lower bound.

In order to appreciate the difference between the fixed-margin and variable-margin semantics, we must consider models in which they come apart. In order to do this, let's revise our interpretation of the predicate B. Instead of letting $B(n)$ represent the claim that any man with just n hairs is bald, we let $B(x)$ represent the claim that any man with just x square inches of scalp coverage is bald, where x now ranges over the non-negative reals.

Given that this revision yields a densely ordered series of possible cut-off points for B, let us also revise our statement of Gómez-Torrente's property to accommodate this change.

Gómez-Torrente's Property (Second Version):

> If, at any world w, there are transparent and non-transparent propositions of the form $B(x)$, then for some x and some sufficiently large m we have:
>
> $$w \models K(K^m B(x) \ \& \ \neg K^m B(x + \varepsilon)).$$
>
> That is, $K^m B(x)$ is a predicate whose cut-off point is known at w to occur between x and $x + \varepsilon$. Here, ε is to be some positive

---

[7]In "Vagueness and Margin for Error Principles" (§3.2).

[8]Equivalent in the sense that a sentence A will be true at a world w in a fixed-margin model $\langle W, d, r, \text{VAL} \rangle$ just in case it is true at w in the variable-margin model $\langle W, d, r, \text{VAL}' \rangle$ where VAL and VAL$'$ are alike with respect to their assignments of truth values to atomic sentences.
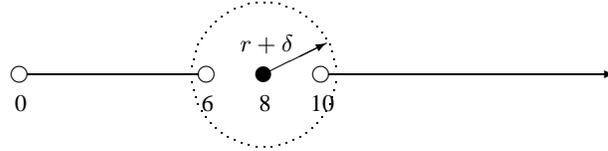
number that is sufficiently small for a boundary between x and $x + \varepsilon$ to count as a sharp cut-off.

This more general version of Gómez-Torrente's property fails to hold in the following model:

Countermodel:

- Let $W = \{s_y : y \in \mathbb{R} \ \& \ (0 < y < 6 \ \text{ or } \ y = 8 \ \text{ or } \ 10 < y)\}$

- For any $s_y \in W$: $s_y \models B(x)$ iff $0 \leq x < y$.

- Let $d(s_y, s_z) = |y - z|$, for all $s_y, s_z \in W$.

- Let $r = 2$.

The structure of the model is represented by the following diagram:

In this model, at s_8 $B(0)$ is transparent, while $B(2)$, for example, is not. Yet still we have the following, in conflict with Gómez-Torrente's property:

$s_8 \not\models K(K^m B(x) \ \& \ \neg K^m B(x + \varepsilon))$,
for any x and m, as long as $\varepsilon \leq 4$.

For suppose $s_8 \models K(K^m B(x) \ \& \ \neg K^m B(x + \varepsilon))$. Then for some $\delta > 0$, $s_{6-\delta} \models K^m B(x)$ and $s_{10+\delta} \models \neg K^m B(x + \varepsilon)$. But since $s_{6-\delta} \models K^m B(x)$, x must be less than 6, since otherwise, $B(x)$ would be false at $s_{6-\delta}$. Also, $(x + \varepsilon)$ must be greater than 10, since whenever $(x + \varepsilon) \leq 10$, $B(x + \varepsilon)$ is transparent at $s_{10+\delta}$. So $\varepsilon > 4$.

We now have a variable-margin model that is not a GT-model. But there is a notably strange feature of the model—the gappiness of its set of possible worlds, in particular, the absence of worlds similar to s_8. Let me introduce some terminology to describe this awkward feature. I'll say that a possible worlds model with a similarity relation on its worlds is a "stepping-stone" model, relative to a degree of similarity r, just in case for any two worlds w and v in the model, there is a

11

non-circuitous sequence of "stepping-stone" worlds "between" w and v, each within a degree of simliarity r from the next:

- For any $w, v \in W$, if $d(w, v) \leq n \cdot r$ (where $n \geq 2$), then:

$$(\exists u_1 \cdots u_{n-1} \in W)(d(w, u_1) \leq r \;\&\; d(u_1, u_2) \leq r \;\&\; \cdots$$
$$\&\; d(u_{n-1}, v) \leq r)$$

What's notably strange about the model that lacks Gómez-Torrente's property is that it is not a stepping-stone model. In fact all models that don't have Gómez-Torrente's property are not stepping-stone models. Every stepping-stone model is a GT-model. To see that this is so, we should first note that although the **B** schema is not validated in all stepping-models given the variable-margin semantics, a close relative of it is:

$$\mathbf{B}^{+}: \; \models \neg A \longrightarrow K\neg KKA.$$

In fact, for the validity of $\mathbf{B}^{+}$ all we need is a weak stepping-stone assumption—that whenever two worlds w and v are in within *twice* the margin for error r, there is a world u "between" them within the margin for error of each:

- $(\forall w, v \in W)(d(w, v) \leq 2r \longrightarrow (\exists u \in W)(d(w, u) \leq r \;\&\; d(u, v) \leq r))$

    - For now suppose that $w \models \neg A$.
    - Then $u \models \neg KA$, for any u such that $d(w, u) \leq r$, by VAL K.
    - Thus $v \models \neg KKA$, for any v such that $d(w, v) \leq 2r$, by existence of a stepping stone u, and VAL K.
    - Thus $w \models K\neg KKA$ by VAL K (since $2r > r$).

Now to show that, given the variable margin semantics, all stepping-stone models are GT-models.

- Suppose that $B(x)$ is transparent at $w \in W$ but that $B(x + \varepsilon)$ is not transparent at w. Here x and $x + \varepsilon$ are to be two things *very close* together in a sorites space (series, line, plane, solid, etc.) for B.

- That is, we have:

  $w \models K^l B(x)$, all l;

  $w \models \neg K^m B(x + \varepsilon)$, some sufficiently large m.

12

- Now by $\mathbf{B}^+$, $w \models K \neg K^{m+2} B(x + \varepsilon)$.

- But $w \models K K^{m+2} B(x)$, since $B(x)$ is transparent at w.

- So, given that $(KA \ \& \ KB)$ implies $K(A \ \& \ B)$, we get:

  $w \models K(K^{m+2} B(x) \ \& \ \neg K^{m+2} B(x + \varepsilon))$.

  That is, $K^{m+2} B(x)$ is a predicate whose cut-off is known at w to occur between x and $x + \varepsilon$.

It seems plausible that we *should* restrict our attention just to stepping-stone models, otherwise knowledge of simple sentences—ones not themselves containing K—would be too easy to come by.

3   Epistemicist Options

Now that we have convinced ourselves that Gómez-Torrente's property is sufficiently widespread to create a worry, let's state the difficulty the result poses for Williamson. The result is that if some but not all propositions of the form $B(x)$ are transparent, then for some sufficiently large m, $K^m B(x)$ is a predicate whose cut-off point is known. The difficulty for Williamson is that *if* this predicate is also a vague one, then there is a vague predicate for which his account is false. Two options for Williamson immediately present themselves: (i) deny the existence of the transparent propositions, or (ii) deny that $K^m B(x)$ is vague.

Is option (ii) feasible? In fact, why think that $K^m B(x)$ is a vague predicate at all? How one goes about addressing this question will depend on what exactly one thinks vagueness amounts to. One thought about the defining feature of vague predicates is that they have *borderline cases*. Does $K^m B(x)$ have borderline cases, where m is a fairly large number? I find it difficult to even tackle this question, since when m is large, even not very large—say as large as ten—I find that I don't even know what criteria I would use to assess whether $K^m B(x)$ applied to a given number. I don't know how to find out whether I know that I know that I know that I know that I know that such and such holds. But to conclude that $K^m B(x)$ has borderline cases I would have to apply criteria for deciding whether the predicate applied in a given case, and then find that even after application of my criteria, I could not decide the question. If I have no such criteria, then I am not in a position to decide that $K^m B(x)$ does have borderline cases.

13

Coming at the question from a different angle, we might ask whether $K^m B(x)$ has another feature characteristic of vague predicates: would a sorites argument involving it hold any appeal for us? In particular, would we be inclined to accept a *sorites premise* for this predicate—namely, $(\forall n)(K^m B(n) \longrightarrow K^m B(n+1))$? Again, since I don't know how to assess whether I know that I know that I know that I know that I know that I know that such and such holds, I can't say that this particular sorites premise holds much sway over me.

Nevertheless, it should be noted that Williamson along with other epistemicists takes the knowledge operator to give content to the 'determinately' operator used to express vagueness. Thus for an epistemicist, the question whether $K^m B(x)$ is vague is just the question whether $\Delta^m B(x)$ is vague, where Δ is the 'determinately' operator. If this predicate is not vague, then presumably the predicate B is not itself indefinitely higher-order vague—which would be an interesting consequence of the existence of transparent propositions.

But then, with regard to option (i)—the option of denying the existence of transparent propositions—if the predicate $K^m B(x)$ is to be understood as giving content to the predicate $\Delta^m B(x)$, then the question whether there are transparent propositions stands or falls with the question whether there are absolutely determinate cases for the predicate B, with the question, for instance, of whether it is not only determinate that any man with zero hairs is bald, but whether it is also determinate that it is determinate, and determinate that it is determinate that it is determinate, and so on. That vague predicates such as 'bald' have absolutely determinate cases is something few anti-epistemicists would be willing to reject, and so an anti-epistemicist might think that if the results of this paper require us to give up the existence of transparent propositions, then that shows us precisely that the knowledge operator does not have the same content as the determinately operator.

We should also at this point discuss the fact (mentioned earlier in nt. 2), that while the knowledge operator K used in stating Gómez-Torrente's property has here been interpreted as meaning 'it is known that', Gómez-Torrente himself gives the operator the weaker interpretation 'it is know*able* that'. We may say that a proposition p is *translucent* just in case p is knowable, and it is knowable that p is knowable, and it is knowable that it is knowable that p is knowable, and so on, for any number of iterations of the weaker knowledge operator 'it is knowable that'. If the anti-epistemicist consequence follows from the existence not just of transparent propositions but also of translucent ones,

then option (ii)—the option of denying the vagueness of $K^m B(x)$—
seems the only plausible way of upholding Williamson's epistemicism,
since we would be hard pressed to deny that $B(0)$, for example, is
translucent. What we need to question, then, is whether the formal
features of the meaning of 'it is known that' that were appealed to in
proving our generalizations of Gómez-Torrente's result hold equally
well for the operator 'it is knowable that'. If not, then some different
argument for the new, stronger result would be required.

One might claim that, on a margin for error epistemology like
Williamson's, the difference between 'it is known that' and 'it is know-
able that' should amount to nothing more than a difference in how
much of a margin for error is required—that knowability should sim-
ply require less of a margin than actual knowledge.[9] If this claim could
be upheld, then since nothing in our arguments assumed any specific
margin for error, those arguments would work equally well here.

Alternatively, one might deny that *knowability* ever requires a mar-
gin for error of at least some given amount, and propose instead that
knowability requires just some positive margin for error or other. To
capture this conception, letting $K_\diamond$ abbreviate 'it is knowable that', we
might adopt a truth clause like the following:

- VAL $K_\diamond$: $w \models K_\diamond A$ iff $(\exists r > 0)(\forall v \in W)(d(w,v) \leq r \longrightarrow v \models A)$.[10]

But this semantics for the knowability operator, given a certain natural
assumption about the relevant class of models, has an even more trou-
bling anti-epistemicist consequence, namely, that if the cut-off point
for the predicate B (for example) is located within a region, then it is
knowable that the cut-off is located in that region, however narrow the
region might be.

The assumption in question is that we have a model in which the
similarity measure on worlds tracks how similar they are with respect
to their cut-off point for the predicate B. (The cropped models dis-
cussed earlier all had this property.) In other words, we have a model
$\langle W, d, \text{VAL} \rangle$ in which:

---

[9] Just as clarity ('it is clearly the case that') should require more of a margin than
actual knowledge. See nt. 3.

[10] We require the margin r to be positive, since otherwise knowability would co-
incide exactly with truth, which on Williamson's view, it does not. We should
note that the proposed clause for $K_\diamond$ yields the validity of the following schema:
$K_\diamond A \longrightarrow K_\diamond K_\diamond A$. The proof depends on the fact that d satisfies the triangle
inequality.

- $d(w,v) = |\text{least upper bound}(S_w) - \text{least upper bound}(S_v)|$.

  Here for each world $u \in W$, S_u is the set $\{x : u \models B(x)\}$, and every set S_u is assumed to be closed under *less than* on the non-negative reals.

So, for example, if in a model we have that at a world w the least number of square inches of scalp coverage one could have without being bald is 15, and that at a world v the least number of square inches of scalp coverage one could have without being bald is 15.2, then in the model $d(w,v) = .2$.

At any world in such a model, if the cut-off point c for B is located within some extended interval $[c - \varepsilon, c + \varepsilon]$ it is knowable that the cut-off point is located within that interval, *however narrow that interval might be*.

- For suppose we have a model in which c is the cut-off for B at a world w, in the sense that c is the least upper bound of the set S_w.

- Now suppose that $d(w,v) \leq \varepsilon/2$, where $0 < \varepsilon \leq c$.

  Then $|\text{lub}(S_w) - \text{lub}(S_v)| \leq \varepsilon/2$. So $c - \varepsilon/2 \leq \text{lub}(S_v) \leq c + \varepsilon/2$.

  - Then $c - \varepsilon < \text{lub}(S_v)$. So $c - \varepsilon \in S_v$ and $v \models B(c - \varepsilon)$.
  - Also, $\text{lub}(S_v) < c + \varepsilon$. So $c + \varepsilon \notin S_v$ and $v \models \neg B(c + \varepsilon)$.
  - So $v \models B(c - \varepsilon) \,\&\, \neg B(c + \varepsilon)$.
  - So $w \models K_\diamond(B(c - \varepsilon) \,\&\, \neg B(c + \varepsilon))$, (no matter how small ε may be), since v was arbitrary and $\varepsilon/2 > 0$.

Thus the adoption of a margin for error semantics for *knowability* will have one anti-epistemicst consequence if it is required that the margin for error be always at least some given positive amount, and will have another anti-epistemicist consequence if it is required only that the margin for error be just some positive amount or other.

To sum up, in order to avoid the difficulties raised in this paper, the proponent of Williamson's epistemicism and margin for error semantics for knowledge must:

1. (a) Keep the fixed-margin semantics for actual knowledge, but deny either the existence of transparent propositions or that $K^m B(x)$ is vague for large m; or

(b) Keep the variable-margin semantics for actual knowledge, but either deny the existence of transparent propositions, deny that $K^m B(x)$ is vague for large m, or deny that we should restrict our attention just to stepping-stone models;

2. (a) Keep the positive-minimum margin for error semantics for knowability, but either deny the existence of *translucent* propositions, deny that $K^m B(x)$ is vague for large m, or deny that we should restrict our attention just to stepping-stone models; or

(b) Keep the no-minimum margin for error semantics for knowability, but either deny the existence of *translucent* propositions, deny that $K^m B(x)$ is vague for large m, or deny that we should restrict our attention just to models in which the similarity measure tracks distances between cut-off points of the vague predicate in question.

Among these, there is no completely unproblematic option. Which among them is the best, I'll leave as an open question.

References

Gómez-Torrente, M. (1997), 'Two Problems for an Epistemicist View of Vagueness', in E. Villanueva, ed., *Philosophical Issues, 8: Truth*, Ridgeview, Atascadero, CA, pp. 237–245.

Gómez-Torrente, M. (n.d.), Vagueness and Margin for Error Principles. Forthcoming.

Williamson, T. (1994), *Vagueness*, Routledge, London.

Williamson, T. (1997), 'Replies to Commentators', in E. Villanueva, ed., *Philosophical Issues, 8: Truth*, Ridgeview, Atascadero, CA, pp. 255–265.