




Supervaluationism about Vague Names Cannot Account for Statements about Those Names

Hugo Heagren

ABSTRACT Vague names, like “Everest” and “Belle Epoque” seem to refer to objects without clear boundaries. Supervaluationism claims that this vagueness is a feature of language, not of the objects referred to; vagueness in names is just ambiguity between many possible referents. This general idea admits of two more specific versions. Both give similar treatments of standard uses of vague names, but have very different results for other cases, such as reference achieved by descriptions including mentioned names. Considering two examples, I show neither variant of supervaluationism can account for the truth of all types of sentences about those names themselves. If I am right that these types exhaust supervaluationism, the theory is shown to be false. This problem closely resembles others in the supervaluationist literature about disquotation failure for truth. Treatments of vague truth and vague reference come apart though, and I show that the two problems are different enough that none of the popular solutions will succeed for reference. I consider—and reject—several specific objections, and two more general ways to recover supervaluationism following my arguments. I conclude that supervaluationism is at best a useful formalism for some kinds of vagueness reference, but fails as a general account.

KEYWORDS reference; semantics; supervaluationism; vagueness

Hugo Heagren, King’s College London, United Kingdom  hugo.heagren@kcl.ac.uk  0009-0005-6052-3369

©  *FORUM PHILOSOPHICUM* 29 (2024) no. 2, 317–34
ISSN 1426-1898 E-ISSN 2353-7043

SUBM. 22 December 2023 Acc. 28 May 2024
DOI:10.35765/forphil.2024.2902.05

Some proper names are vague, in that there is some sort of indeterminacy in what they refer to. Examples include “(The) Thames,” “Everest” and “Belle Epoque.” It is at least not obvious where (or when) the referents of such names begin and end. Views on the nature of this indeterminacy fall into two camps. Some (Tye 1990) think that the objects themselves (the mountain Everest, the river Thames, etc.) are really ontologically “fuzzy”—they simply do not have sharp boundaries. Others think that it is the *names* which are vague—there are no fuzzy objects, but the language we use to refer to some objects does not capture this: it “refers vaguely.”

Supervaluationism is a popular account of the latter type, which analyses vagueness as ambiguity between precise interpretations (vague predication as ambiguity between precise predicates, vague reference as ambiguity between precise referents, etc.). Unfortunately, this makes supervaluationism incapable of correctly analysing sentences with truth conditions which *depend* on vagueness, as I will show. Since such sentences exist (even appearing in the supervaluationist literature!) I conclude that supervaluationism is an inadequate account of vague proper names. A corollary of this conclusion is that supervaluationism fails as an account of vagueness in general, since it cannot analyse all vague phenomena.

In the first section, I describe how supervaluationism accounts for vague names and sentences containing them, and distinguish two more specific types of supervaluationist theory. In the second section I show that neither type of theory can account for both of two exemplar sentences. Given that these two types exhaust the possible supervaluationist theories of names, I conclude that no such theory is correct. I consider various objections, and finally the resulting prospects for supervaluationism.

SUPERVALUATIONISM

Supervaluationism is an account of vagueness which can in principle be applied to any vague phenomena. The objection in the body of this paper is concerned only with supervaluationist accounts of vague proper names, and the objects they refer to. The corollaries drawn afterwards concern it as a theory of vagueness in general. Supervaluationism has two motivating ideas: that there are no truly vague objects (any objects which exist are sharply bounded) and that vague meaning in language is nothing more than ambiguity between many possible non-vague meanings.¹

1. On some readings, the second idea entails the first, and thus anyone committed to the latter is committed to the former. All that is relevant for this paper is that supervaluationism (as I understand it here) is in fact committed to both, regardless of their relation.

According to supervaluationism about vague names, every vague name has a (possibly indeterminate) number of non-vague interpretations, between which the term is ambiguous. These interpretations are called “precisifications” or “sharpenings” in the literature (e.g. Varzi 2001).

“Everest” is a vague name, but there are many ways a sharp boundary could be drawn around Everest itself. Each of these different chunks of matter is a precisification of Everest. A sentence containing a vague term is then true (or “supertrue” for the supervaluationist) just if it is true under all precisifications of the term. It is false (“superfalse”) just if it is false under all precisifications. Otherwise its truth value is indeterminate (these are the “borderline” cases).

Consider, for example, the following true statement:

1. Everest is entirely in Asia.

Supervaluationism can explain why 1 is true, even though “Everest” is a vague name. Suppose a series of precisifications of Everest: E1, E2, E3, etc., each a precisely-bounded area of matter around a certain peak in the landscape. A series of precisified sentences will result, by considering Everest as each of these objects in turn:

- 1.1. Everest is entirely in Asia (*where “Everest” refers to E1*).
- 1.2. Everest is entirely in Asia (*where “Everest” refers to E2*).
- 1.3. Everest is entirely in Asia (*where “Everest” refers to E3*).
- 1.4. etc...

Now, since every precisification of Everest is entirely in Asia, every one of these sentences will be true. So for supervaluationism, the original “Everest is entirely in Asia” will be true as well. A similar process would show that “Everest is entirely made of marzipan” is false (since no sharpening of Everest is entirely made of marzipan) and that “Everest is exactly 12 miles wide” is indeterminate (since this will be true of some sharpenings, but not others).

Varzi (2001) gives the clearest and most detailed supervaluationist account of vague names. Others who discuss supervaluationist accounts of vagueness in general mention their application to reference in passing, though without more development. Williamson (1994) mentions the name “Toronto” (145) as an example of vagueness, and seems to have a roughly referential picture of meaning in mind when discussing supervaluationism. Keefe (2000, 159–60) also mentions “Toronto” and gives a brief

supervaluationist account of vague reference (which she compares and prefers to an account of precise reference to vague objects). Lewis (1988) also seems to support a supervaluationist analysis of vague names.

It is important to stress that these are supervaluationist accounts of vague reference. What is supposed to be vague or indeterminate is the relationship between names and their referents, not (necessarily) the truth conditions of sentences containing such names (though these might be derivatively vague). This is in contrast to most presentations of supervaluationism (e.g. Keefe 2000), which operate entirely in terms of truth.

TWO TYPES OF SUPERVALUATIONISM

The basic insight is clear enough, but full formalisations of supervaluationist semantics vary enormously (see, for an overview, Varzi 2007). For my purposes, only one difference between theories is relevant: the interpretation of tokens of vague names which are not used (that is, those which are mentioned, or tokened in some other way, such as by spelling them out).

All supervaluationist theories agree that for a name to be vague is for its reference to be ambiguous between many precise referents with which it could be interpreted. And all supervaluationist theories proceed by analysing each sentence under a number of different precisifications. Under each precisification, uses of the relevant vague name are interpreted as referring to some precise object.

On some such semantics, when a sentence with a vague name is analysed, the name is always interpreted as having the relevant precise object as its referent, whether the name is used or tokened in some other way—the name's semantics change entirely on each precisification. On theories like this, under each precisification, a description like "the referent of 'Everest'" always refers to a single precise object (though a different one under each precisification). As such, on theories like this, a name is vague (i.e. ambiguous) before the supervaluationist analysis is applied, but on each precisification it is not vague/ambiguous—ambiguity is understood as the availability of many different precisifications, but on each precisification one such precisification has been chosen, so the name is not ambiguous (i.e. vague). Call these A theories.

On other theories, the name remains vague on each precisification. That is, under any precisification, it is still true that the name refers ambiguously to many objects, but only one of these is considered under each precisification. The name itself is interpreted as still having ambiguous reference. On theories like this, the referent of a description like "the referent of 'Everest'" would either be nothing (like "the present King of France"),

or indeterminate, or ambiguous between various precise objects, though “Everest” (under each precisification) will be interpreted as referring to some particular precise object. Call these B theories.

To be clear: A and B theories agree that for a name to be vague simpliciter (not on any precisification) is for there to be many ways of making its referent precise. The point of contention is how names behave under each precisification. A theories say no name is vague on any precisification of that name. B theories say the opposite.

It is not necessary to go into detail here about the theoretical motivations for each type of theory, but it should be clear that both types can correctly analyse sentences such as 1.²

THE OBJECTION

Consider the following sentence:

2. The referent or referents of “Everest” is/are identical with Everest.

I take it that 2 is true, as would be any sentence formed by replacing both tokens of “Everest” in 2 with other referring names (vague or otherwise). And I take it that 2 is true because, in general, a used name *n*, and a description of the form “the referent or referents of *n*,” refer to the same thing (at least for referring names *n*, outside of modal or attitudinal contexts). This is an analogue for reference of the T-schema for truth (i.e. *p* iff “*p*” is true) and is so widely accepted that it forms the basis of deflationary theories of reference and “austere” theories of meaning (see e.g. Horwich 1988, 118; McDowell 1977; Båve 2008; Sainsbury 2005, 73). On an A theory, under any precisification, “Everest” will refer to exactly one precise object, though a different object on each precisification. On one precisification the name refers only to E1, on another only E2, and so on. On each precisification, the definite description “The referent or referents of ‘Everest’” will refer to the same object as “Everest” does (since on each precisification, “Everest” only has one referent). Thus, on each precisification, the identity claim in 2 will be true. Since it is true on all precisifications, it will be supertrue. But then an A theory will show that the following is false:

3. Everest’s name is vague.

2. Indeed, the two types are entirely isomorphic, at least for simple sentences like 1 (Varzi 2007, fn.5).

“Everest’s name” is a definite description which refers to “Everest.” On an A theory, under any precisification, “Everest” is interpreted as referring to a single precise object (though again, a different precise object on each precisification).

Yet on any supervaluationist theory, for a name to be vague just is for that name’s reference to be ambiguous between many precise objects. And on an A theory, on each precisification, this is not the case. As above, ambiguity is understood as the availability of many different precisifications, but *on each precisification* one such precisification has been chosen, so the name is not ambiguous (i.e. vague). So on any precisification, it will be false that Everest’s name is vague. Since it is false on all precisifications, the original sentence will be false overall (superfalse). The A theorist then is committed to the falsity of “Everest’s name is vague,” but this is absurd—clearly “Everest” *is* a vague name.

So A theories can analyse sentences such as 2 correctly, but cannot handle sentences like 3.

The situation for B theories is the reverse. On a B theory, under each precisification, only *uses* of a vague name are interpreted as referring to the relevant precise object. Other tokens of the name are interpreted as having ambiguous reference (that is, on a B theory, a sentence like “‘Everest’ has only one referent” is not true on any precisification).

Thus a B theory will correctly analyse 3 as true; “Everest’s name” is a definite description referring to the name “Everest.” And on a B theory, on each precisification, the reference of “Everest” is indeterminate or ambiguous between various precise objects. According to any supervaluationist theory, this is just what it is to be a vague name, so 3 is true.

But this comes at the cost of analysing 2 incorrectly. On a B theory, under any precisification, a used vague name is interpreted as referring to the relevant precise object, but other tokens of the name are not. (This must be true if 3 is to be analysed as true.) Thus, under any precisification, the reference of the description “The referent or referents of ‘Everest’” is indeterminate or ambiguous, (or perhaps if precise, it is a set or plurality of admissible precisifications). But the reference of “Everest,” as used at the end of the sentence, is precise (it will be the relevant precise object E1, or E2, or E3, etc.). The sentence claims that the two are identical. But this cannot be the case—one refers to a single precise object, and the other refers indeterminately or ambiguously. It is not even clear how such an identity claim is to be understood, since it is not obvious *what* the referent or referents of an indeterminately referring description would be.

So B theories can analyse sentences such as 3 correctly, but cannot handle sentences like 2.

The argument against supervaluationism about vague names is then simple. Both 2 and 3 are true. All supervaluationist theories are either A theories or B theories. Neither type can account for the truth of 2 and 3, so neither type is a correct theory, and so no supervaluationist theory of vague names is correct.

I turn now to consider some possible responses to this argument.

RESPONSES

2 is true, but pragmatically says something non-obvious

A supervaluationist might accept that 2 is true, and perhaps even that on a strictly literal interpretation, the supervaluationist analysis is false, *but* nevertheless claim that this is not a problem, because in ordinary parlance sentences like 2 are not used or understood in this strict way—they are elliptical. They might claim, say, that a sentence like 2 is usually understood as meaning something like “every precisification of Everest is one of the objects which ‘Everest’ (ambiguously) refers to.”

A respondent might even go further, and say that the discrepancy between the strict falsity and ordinary-use truth of 2 is just what makes vagueness philosophically interesting, and this is exactly why we need a philosophical treatment of it.

If this is true, then the supervaluationist can accept that my argument above is entirely correct, and still resist its being an objection to their position. There are two significant problems with this view.

First, it seems unlikely. While sentences in ordinary parlance outside of philosophy are often elliptical in the way proposed, “ordinary parlance” about vagueness and reference (and especially about *vague reference*) is almost never outside of philosophy. The New Yorker does not publish bulletins on recent developments in metasemantics. But within philosophical discourse—including that about reference and vagueness—it is reasonable to expect linguistic standards to be higher, and ellipsis of the kind suggested by the respondent is avoided as much as possible, so the strictly literal interpretation is the right one.

I constructed 2 to be as clearly as possible a claim of *identity* between two things, not a claim that one thing is included in a plurality. If the supervaluationist response was right, then much language in philosophy, formulated specifically to be as explicit as possible, does not really mean

what we think it means. This seems unlikely, and the supervaluationist must provide independent motivation and evidence for such a claim.

Second, even if this response is correct, the objection itself still stands, as it will be possible to restate it in other terms. Sentence 2 seems to be a good way of demonstrating it, but the point is that, in general, a used name *n*, and a description of the form “the referent of *n*,” should refer to the same thing (at least for referring names *n*, outside of modal or attitudinal contexts). I intended the discussion of 2 as a good way of demonstrating how this principle fails for supervaluationist accounts of names. But if 2 is in fact used in a non-obvious way to say something else, all that is shown is that my example was badly chosen, not that the objection itself fails.

To respond properly to my objection, the supervaluationist must show that the principle I rely on is actually false—that in general we should not always expect names *n* and definite descriptions “the referent or referents of ‘*n*’” to co-refer (at least, for referring names *n*). A full argument for this is beyond this paper, but I would certainly be surprised to find it was false.

The description “the referent or referents of ‘Everest’” is not apt

A friend of B theories might claim that, since the reference of “Everest” is indeterminate (and this is still true if the name is mentioned, even under a precisification), a description like “the referent or referents of ‘Everest’” is not apt to be used in the way it is in 2. They might say such a description is meaningless, or nonsense, or empty, as would be a phrase like “the referent of cheese.” If so, then the sentence in 2 is not true, so a theory of vague names need not account for its truth. I have admitted that a B theory can account for the truth of sentences like 3, so this response would show that a B theory can account for all the sentences it needs to. This is compatible with accepting my argument in 3.1 that sentences like 2 really do claim an identity—the objector will just claim that 2 is in fact a failed attempt to create an instance of this form, because the definite description doesn’t actually refer to anything.

A response like this would not show conclusively that supervaluationism is correct, but it would show that my objections to it fail, at least for B theories.

This response fails, however.

Depending on the exact specification of the theory, it might be true that descriptions like “the referent or referents of ‘Everest’” are not apt for use in sentences like 2. But this does not defeat the spirit of my objection. The point of the objection is that it is possible to deploy a name in such a way as to refer to the name’s referent, but without using the name (i.e. by instead

mentioning the name in a description), and that on any reasonable semantic theory, using a name in this roundabout way and using it in the more usual way should specify the same object (again, at least for referring names outside of modal or attitudinal contexts). Consider the sentences: “Matt is coming,” “the person called ‘Matt’ is coming,” “the person with this name is coming” (spoken while pointing to a written name the speaker cannot/will not pronounce), “M-A-T-T is coming,” “the person called ‘M-A-T-T’ is coming.” There is an important sense in which the various deployments of the name all achieve the same thing: reference to Matt, regardless of how the name is deployed. Devices like this are often used in front of very young children to obscure which object is referred to, though which object is referred to remains constant regardless of how the name is deployed.

My objection was meant to show that B theories cannot sustain the constancy of reference under redeployment inside certain descriptions, because on any precisification the description with the mentioned name refers to some set or plurality of admissible precisifications (or nothing at all, or is meaningless), but the use refers to a single precise object (the relevant one for that precisification).

The proposed response is that the way I specified the name’s referent in 2 should not be expected to specify the same object as a use of the name would. (As I described it above, the response is that my proposed description is empty or meaningless.)

Even if this is true though, that does not mean there is *no* way to deploy a name circuitously to specify its referent without using it—just that the way I assumed would work in fact does not.

Perhaps the description could be replaced by “the object or objects to which ‘Everest’ (precisely or vaguely) refers (if it refers at all),” or even, in the supervaluationist’s terminology, “the referent or admissible precisified referents of ‘Everest’.” These surely would specify the relevant set or plurality of precise objects.

Sentences constructed like 2 but using these descriptions would still be true, and B supervaluationist theories would still show they are false as described above, so this response fails.

The A/B distinction is not exhaustive

It might be objected that the distinction between A and B supervaluationist theories is not exhaustive. If so, then, although I have shown above that neither A nor B supervaluationist theories are correct, there might be some further type which can handle all the cases I consider, and which *is* correct.

At least in consideration of supervaluationist theories of the meaning of vague *names*, I argue that the distinction between A and B theories is exhaustive.

First, some clarification. Supervaluationist theories of vague names are theories on which the reference (or other type of meaning) of vague names is indeterminate. There are also other supervaluationist theories of meaning and vagueness, some of which account for vague names, but these are not my concern here (e.g. Fraassen 1966). I am only concerned to show that supervaluationist theories of vague names fail.

Any such theory posits some variant of the following picture: there is something appearing to be a single name in our actual language (such as “Everest”), there are more than one candidate for precise referents of this apparent name, and there is some connection between the name and the candidates. Whatever the connection, it must be such that under each precisification (whatever this is taken to mean), the apparent name is related uniquely to one of the candidate objects, and that (non-modal, non-attitudinal) uses of the name refer to that object (so that sentences like 1 come out as true).

There are two ways to fulfil this condition. First, suppose that the semantics of the name differ entirely between precisifications. Thus, under each precisification, there is a name which refers to the relevant object, fulfilling the requirements of the theory. (Whether the *same name* is being used under each precisification is just a matter of how one individuates names, and has no bearing on this paper.) This is an A theory. Now suppose otherwise—it is still true that, under each precisification, the name relates uniquely to one of the candidate objects, and that uses of the name refer to that object. But it is not true that the semantics of the name differ completely between precisifications. Putting aside trivial variations of an A theory (such as where the referent of the name when quoted is some *other* precise object), the result is that the name itself remains ambiguous. On each precisification, one object is “under consideration,” though it is not the sole referent of the name. This must be the case on any supervaluationist theory which is *not* an A theory. But this just is a B theory—so the A/B type distinction is exhaustive for supervaluationist theories of names.

The (same) theory need not always be applied

Lewis provides a supervaluationist solution to the problem of the many, and considers an objection similar to the one I raise:

Supervaluationism works too well: it stops us from ever stating the problem in the first place. The problem supposedly was that all the many candidates had

equal claim to cathood. But under the supervaluationist rule, that may not be said. For under any one way of making the unmade decision, one candidate is picked as a cat. So under any one way of making the decision, the candidates do not have equal claim. What's true under all ways of making the decision is super-true. So what's super-true, and what we should have said, is that the candidates do not have equal claim. Then what's the problem? And yet the problem was stated. So supervaluationism is mistaken (Lewis 1999, 173).

My objection could be put similarly: the supervaluationist claim supposedly was that all the many candidates had equal claim to be the referent of "Everest." But under the (A theory) supervaluationist rule, that may not be said. For under any one way of making the unmade decision, one candidate is picked as the referent of "Everest." So under any one way of making the decision, the candidates do not have equal claim. What's true under all ways of making the decision is super-true. So what's super-true, and what we should have said, is that the candidates do not have equal claim—i.e. that "Everest" is not vague (and therefore that 3 is false). Then what's the claim? And yet the claim was stated.

Lewis' response is that supervaluationism simply shouldn't be applied across the board:

[supervaluationist analysis] should instead be taken as a defeasible presumption. What defeats it, sometimes, is the cardinal principle of pragmatics: The right way to take what is said, if at all possible, is the way that makes sense of the message. Since the supervaluationist rule would have made hash of our statement of the problem, straightway the rule was suspended (Lewis 1999, 173–4).

Lewis' problem is slightly different to mine, but the spirit of his solution is clear enough: supervaluationist analysis is not to be applied indiscriminately across the board. In a similar vein, a supervaluationist might accept that neither an A nor a B theory can deliver the right analysis for all sentences, but object that not all sentences need be treated in the same way: for vague names used as in sentences like 2, an A theory is true, whereas for those in sentences like 3, a B theory is true. This is a more sophisticated version of Lewis' move: no *particular* supervaluationism need be applied across the board, though *some* supervaluationism is applied in each case.

Such a "mixed" theory—Lewis' suggestion or mine—would deliver the correct results for the sentences I have considered. But the onus is on the supervaluationist to provide independent motivations for such a complex

theory, and to provide a satisfactory way of distinguishing the two (or possibly more) types of case. Whether this is possible is beyond the scope of this paper, but it does not seem likely to me.

There are also more formal reasons that such a theory is unlikely to succeed. Even if there is a satisfactory theory on which different tokens of the vague name have different semantics, it is still possible to construct sentences in which a single token of a vague term is used in both of the ways proposed (using anaphora). Consider:

4. Everest's name is vague, and its (the name's) referent or referents is/are identical with Everest.

The whole sentence is true. Since it is a conjunction, it is true because each conjunct is true.

For the first conjunct ("Everest's name is vague") to be analysed as true, "Everest" must be understood under a B theory. But since the "it" in the second conjunct is anaphoric on this use of "Everest," surely it will be interpreted in the same way, making the second conjunct roughly equivalent to an assertion of 2 under a B theory. This would be false, and thus the whole conjunction false, so not even mixed supervaluationism correctly analyses all sentences.

Equivalence is metalinguistic

Keefe defends a supervaluationist account of vagueness. Though she mentions other vague phenomena, her account is developed entirely in terms of truth:

The supervaluationist semantics are built on assignments of *truth-values* to sentences of the language, which Fine calls specifications (2000, 318; emphasis mine).

Keefe recognises a problem for her theory somewhat similar to the one I describe for vague theories of reference. Truth-based supervaluationism must reject the T-schema:

- T. "p" is true if and only if p.

Instances of T are an expression in the object language of the disquotationality of truth. A supervaluationist about vague truth must deny this:

For, suppose “p” is neither true nor false, and so true at some but not all points; “p’ is true” will then be true at no points. At those points where “p” is true, [the right-to-left conditional] is false, since its antecedent is true and its consequent is false; [the right-to-left-conditional] and so (T) cannot be true simpliciter in those cases (Keefe 2000, 214).

But this is counterintuitive, because T seems right. The problem I raise seems similar: 2 is a sort of disquotation (or maybe dis-description) of reference. Instances of T fail when “p” is a vaguely-true/indeterminate sentence: on any precisification, “p” will have a determinate truth value, but “p’ is true” will not. Similarly, sentences like 2 fail for a B theory when “Everest” (or any other name) is a vague name: on any precisification, “Everest” will be vague, and therefore have multiple or indeterminate referents, but uses of “Everest” do not (i.e. “‘Everest’ refers only to Everest,” or other sentences like 2 will not be true).

Keefe’s solution for vague truth is to read “if and only if” in T metalinguistically. She accepts the strange consequences for object-language instances of T when read with the material biconditional, but argues we should read it instead with an appropriately similar metalinguistic notion: entailment. T thus becomes the claim that for any “p,” “p” is supertrue iff “p’ is true” is supertrue. This is the case for a supervaluationist theory of vague truth, so Keefe retains T with a metalinguistic reading. Perhaps a supervaluationist about vague reference can retain 2 with a similar strategy. It is not clear that they can. Keefe’s move is from a relation expressible in the object-language (implication) to an *appropriately similar* but distinct one in the metalanguage (entailment). Her response succeeds because the two relations are different, and it is plausible because the two relations play appropriately similar roles in their respective languages. But the relation at issue in 2 is identity, and the only metalinguistic analogue of identity *is* identity, which is the same relation however it is expressed.

A staunch defender of this type of analysis might just claim that that this not the case: the metalinguistic analogue of identity (at least in the context of 2) is actually a different relation, which makes 2 true. This position collapses into a version of the position discussed and rejected above (3.1).

Moreover, the reasons Keefe gives in support of her move do not apply here. She argues—plausibly—that the intuitions which underpin T are *only* concerned with cases of clear truth, *not* indeterminacy or falsity:

Consider two more of the kinds of things that are said in support of (T) ...
First, if I am in a position to say that p, then I can also say that “p” is true,

since I know that inserting quotation marks and adding “is true” adds nothing to the claim. And second, what it takes for “TW is thin” to be true is just for TW to be thin, no more and no less ...

Now, thoughts such as these do ensure that it cannot be the case that “p” is true while “‘p’ is true” is anything other than true... . It is less clear what follows about the middle case in which “p” is indeterminate. Neither the indeterminacy nor the falsity of “‘p’ is true” would violate the claim that it takes no more for “TW is thin” to be true than that TW is thin, for we are not dealing here with the situation in which TW is thin. Similarly, if “p” is indeterminate I cannot be in the position to assert p and go on to add quotation marks and “is true,” so the first justification dictates nothing about that scenario either (Keefe 2000, 216).

Keefe uses this discrepancy between our strong intuitions for the good cases, and lack of intuitions for the others, to argue for her metalinguistic reading of T. Since our intuitions are silent on what can be derived from claims like “‘p’ is indeterminate,” a correct reading of T should be similarly silent, as Keefe’s metalinguistic reading of T is.

So Keefe adduces two types of evidence. The first is epistemic: whenever I am in a position to say that p, I am similarly in a position to say that “p” is true. This supports the right-to-left direction of T. Keefe isn’t explicit, but presumably this support goes the other way too: whenever I am in a position to say that “p” is true, I am similarly in a position to say that p. This would support the left-to-right direction of T. Furthermore, Keefe claims we *don’t* have these kinds of intuitions for other cases: claims that “p” is false, or that “p” is indeterminate. The second kind of evidence is alethic or metaphysical: “what it takes for ‘TW is thin’ to be true is just for TW to be thin, no more and no less.”

Both kinds of evidence have analogues for reference. First, when I am in a position to use the name “Everest,” I am similarly in a position to use the definite description “the referent of referents of ‘Everest.’” Second, just as we can move by entailment between the claims that TW is thin and that “TW is thin” is true, so we can move by identity between Everest and the referent of “Everest.” Both of these intuitions are common in the literature, and often marshalled by deflationary theorists of reference (e.g. Båve 2008; Horwich 1988, 118 is especially similar to my examples), and “austere” theories of the meaning of names (McDowell 1977; Sainsbury 2005, see esp. 73).

But unlike the intuitions Keefe considers, these referential counterparts are *not* limited to the “good cases” (i.e. not limited to non-vague names). Indeed,

appeals to exactly such intuitions in other literature are often treated as an effective way of dealing with the less clear cases of reference. Sainsbury's entire (2005) project of recovering reference without referents relies on the utility of sentences like "Vulcan' refers to Vulcan" (where "Vulcan" is an empty name) and cases of vague names are never distinguished from non-vague ones in Sainsbury's work or (for example) Horwich's (1988). Keefe's motivation required a discrepancy between the good and the bad cases, but there is no such discrepancy in this analogous evidence. So although the evidence Keefe offers for her solution does have an analogue in reference, it would not motivate an analogous solution, even if one were available.

Semantics are precisification-specific

Williamson (1994) also recognises a problem of disquotation. His diagnosis is that we usually think truth is disquotational from the object-language to the metalanguage (i.e. "p" is true iff p), but that if it is then bivalence holds. The supervaluationist about truth must reject bivalence (because there are indeterminate sentences), so supervaluationism about truth is inconsistent with our intuitions.

Ultimately, Williamson argues against supervaluationism, but he suggests a simple fix for the supervaluationist to avoid the problem of disquotation. Add quotes and a predicate "true_T" to the object language. The value of truth_T for a quoted sentence on a precisification is set to correspond to the truth value of that quoted sentence on that precisification. Finally, we identify truth (the truth that philosophers study, and that can be vague) with truth_T.

An epistemic interpretation of supervaluationism emerges. The sentence "he is bald" (said of a clearly bald man) is true (and therefore true_T) on every precisification. The notion of supertruth drops away, and is functionally replaced with knowledge: it is possible to *know* that he is bald (or that "he is bald" is true), because whichever precisification is chosen, the sentence is true. Falsehood behaves similarly. The advantage of this approach is that truth_T is disquotational, but still palatable for the supervaluationist: "The supervaluationist allows that either 'A' is true_T or 'Not A' is true_T, for this is to allow no more than that either A or not A. In Fine's phrase, the vagueness of 'trues_T' [sic] waxes and wanes with the vagueness of the given sentence" (Williamson 1994, 163). Like Keefe, Williamson's discussion is couched entirely in terms of truth, though a supervaluationist theory of reference could be developed along similar lines. The idea is to abandon the substantive "super" part of supervaluationism, but retain the insight that vagueness can be analysed by comparing the various ways a vague case could be made precise—the *mere* values. The mere values in the referential case will be the precise chunks

of matter which are admissible candidates for Everest. 2 and 3 are already written in a language with names and quotes (and identity, required where Williamson's examples use a biconditional). Instead of Williamson's truth_T , consider reference_T, where the referent_T of a name on a precisification is the precise referent of that name on that precisification.

On this model, both 2 and 3 come out true. "The referent or referents of 'Everest'" is interpreted as (equivalent to) "the referent_T of 'Everest'," which will be identical to the referent_T of "Everest."

3 says that Everest's name is vague. A reasonable paraphrase of a claim of vague-name-hood is that a name does not refer_T to the same object on every precisification: i.e. that there is no object such that on every precisification, that is the referent_T of "Everest." This will also come out true on the proposed account. So 3 is true too. This is a Pyrrhic victory. The referential equivalent of Williamson's proposed supervaluationism is formally adequate, but it sacrifices the thing which the supervaluationist wants: an analysis of vague reference.

There are two interpretations of Williamson's picture. According to Williamson's own, in terms of truth, "of any admissible valuation, we can ask whether it assigns truth to all and only the true sentences of the language and falsity to all and only the false ones. At most one valuation has that property" (1994, 164). Williamson's valuations are assignments of truth values to sentences. On the views discussed in this paper, precisifications are assignments of a precise referent to a term. On the analogous interpretation for reference, then, for each vague name there is exactly one correct and precise object, and we can ask which precisification assigns that object as the referent of that name.

Williamson goes on: "But then any other valuation will assign truth-values incorrectly, so how can it be admissible?" Similarly for reference: if only one precisification is "right," how could the others be admissible? The driving idea of supervaluationism was that the reference of a vague name is ambiguity of reference writ large. But if there is one correct precisification, then the reference is not *ambiguous*, it's just *unknown*. On this interpretation, then—by the supervaluationist's own lights—names like "Everest" are not really vague, and the supervaluationist has nothing to analyse. Williamson himself makes an analogous objection to supervaluationism about truth.

There is an alternative interpretation of Williamson's suggestion which he does not consider, but which might seem more palatable to a supervaluationist about reference. On this reading, there is no one correct precisification, and the existence of multiple admissible precisifications is taken as an analysis of vague reference. Once again, vagueness is ambiguity writ large, and this interpretation delivers actual ambiguity, and so actual vagueness, contra Williamson's reading.

But again, it is not clear what the object of analysis is. Reference is identified with reference_p, and reference_t is only defined per-precisification. On this reading, sentences like 2 and 3 come out true, but there is no global notion of *vague reference* for the supervaluationist to analyse: there is no way to complete the sentence “vague reference is like precise reference, except that ..., whereas precise reference” On this interpretation, supervaluationism is not an analysis of *vague reference*. On either reading, Williamson’s fix saves supervaluationism in name only, not as a substantive analysis of vague names.

PROSPECTS FOR SUPERVALUATIONISM

This argument demonstrates that supervaluationism fails as an account of vague reference. Similar arguments are made against supervaluationist theories of vague truth and predication—though as I have demonstrated, the responses made to those objections are not applicable to mine, nor are there obvious applicable counterparts to these responses which would succeed for reference where those responses do for truth. Supervaluationism is at best a successful analysis of *some* vague phenomena, but not all. It therefore fails as an account of vagueness in general. But all is not lost. I see two significant options still open to a defender of supervaluationism.

First, they might claim that there simply is no vague reference: all vagueness is linguistic, but not referential. This makes my objection entirely irrelevant. This line might naturally continue that not only is all vagueness linguistic, but it is also all *alethic*; the only vague phenomenon is truth. A supervaluationist theory of truth can then take on the sentences above, and treat my objection with one of the responses I have discussed.

This seems an unlikely line given that so many of the defenders of supervaluationism seem entirely comfortable extending the approach to all vague phenomena (though their theories are in fact developed entirely in terms of truth).

Alternatively, a supervaluationist could alter their methodological stance and say that supervaluationism should not be taken as an account of the *nature* of vagueness, but rather as a framework for *formalisation* of vague phenomena. They could claim that—although the framework does not entirely reflect all vague phenomena, it accurately reflects most central cases, and only falls down in unusual scenarios like sentences which discuss vagueness.

This would leave supervaluationism in a similar position to things like possible world semantics or certain formal theories of truth: everyone agrees that these are useful—even elegant—ways of formalising certain questions, but whether this fact has any metaphysical import is hotly debated, and often doubted. (The frequent use of possible worlds in formal presentations

of supervaluationism only strengthens this analogy.) A position like this might be one interpretation of the “pragmatic rule” response made by Lewis (1999). Like Lewis’ position, this account would fail to analyse philosophical discourse about vagueness.

This seems to me the most promising way of looking at supervaluationism—as a useful but limited framework. Supervaluationist accounts either fail to cover all cases (A and B theories), or fail in their role as analyses of vague reference (the Williamsonian reference_T). The theory can thus be taken as a useful formal tool for thinking about vagueness, though not as a substantive account. Of course, such a theory will only remain useful until a correct (or better) theory of vagueness is developed and formalised, though this does not seem a particularly pressing concern.

CONCLUSION

To conclude, supervaluationism accounts for vague names as semantically ambiguous between many, non-vague meanings. A full semantics for such a theory could analyse each instance of a vague name in either of two ways (or analogously with Williamson’s account of vague truth), but none of these can consistently provide a correct analysis. Supervaluationism cannot be a correct theory of vague names and is at best a useful and defeasible formalism for work on vagueness.

BIBLIOGRAPHY

- Båve, Arvid. 2008. “A deflationary theory of reference.” *Synthese* 169 (1): 51–73.
- Fraassen, Bas van. 1966. “Singular Terms, Truth-Value Gaps, and Free Logic.” *The Journal of Philosophy* 62 (17): 481–95. <https://doi.org/10.2307/2024549>.
- Horwich, Paul. 1988. *Meaning*. Oxford: Oxford University Press.
- Keefe, Rosanna. 2000. *Theories of Vagueness*. Cambridge: Cambridge University Press.
- Lewis, David. 1988. “Vague identity: Evans misunderstood.” *Analysis* 48 (3). <https://doi.org/10.1093/analys/48.3.128>.
- . 1999. “Many, but almost one.” In *Papers in Metaphysics and Epistemology*. Cambridge: Cambridge University Press.
- McDowell, John. 1977. “On the Sense and Reference of a Proper Name.” *Mind* LXXXVI (342): 159–85. <https://doi.org/10.1093/mind/lxxxvi.342.159>.
- Sainsbury, Mark. 2005. *Reference Without Referents*. Oxford: Oxford University Press.
- Tye, Michael. 1990. “Vague Objects.” *Mind* XCIX (396): 535–57. <https://doi.org/10.1093/mind/XCIX.396.535>.
- Varzi, Achille C. 2001. “Vagueness in geography.” *Philosophy & Geography* 4 (1): 49–65. <https://doi.org/10.1080/10903770124125>.
- . 2007. “Supervaluationism and Its Logics.” *Mind* 116 (463): 633–76. <https://doi.org/10.1093/mind/fzm633>.
- Williamson, Timothy. 1994. *Vagueness*. London: Routledge.