

# What Can't Be Evaluated, Can't Be Evaluated; and It Can't Be Supervalued Either

Jerry Fodor and Ernie Lepore

*The worry.*

This is a long paper with a long title, but its moral is succinct. There are supposed to be two, closely related, philosophical problems about sentences<sup>1</sup> with truth value gaps: *If a sentence can't be semantically evaluated, how can it mean anything at all?* and *How can classical logic be preserved for a language which contains such sentences?* We are neutral on whether either of these supposed problems is real. But we claim that, if either is, supervaluation won't solve it.

There are lots of sentences that contain empty names; there are lots of sentences whose presuppositions fail; and there are lots of sentences whose predicates are vague (by some author's lights -- Kamp and Partee (ms) -- all predicates are). It seems plausible that at least some of these sentences are incapable of semantic evaluation; a fortiori, that they lack truth values.<sup>2</sup> Let's suppose, for a working example, that 'the present king of France is bald' is one of these.

Now, it seems that such sentences can't simply be meaningless since complex sentences of which they are constituents can be logically true or logically false. For example, 'either it is the case that the present king of France is bald or it is not the case that the present king of France is bald' looks to express a logical truth, hence, a fortiori, to be truth valuable; hence, a fortiori, to be meaningful. Surely, the constituents of a meaningful sentence can't themselves be meaningless? So, the first of our two philosophical problems has the form of a dilemma: On the one hand, how can a sentence be meaningful if it can't be semantically evaluated? But, on the other hand, how can a meaningless sentence be a constituent of a sentence that has a truth value?

Here's the second problem: Philosophers worry that once bivalence fails and truth-value gaps are introduced into semantics, classical logic will have to go. In particular, for the sorts of reasons we've just been considering, they worry about the law of excluded middle. Suppose P is a sentence that lacks truth value. Then, presumably, so too is Not-P; and so too is  $[P \vee \text{Not-P}]$ . We confess to not losing sleep over this, but there are purists who can't bear it.

You might suppose that a philosopher who has taken these worries seriously is just going to have to insist that there are no sentences with truth-value gaps after all. However, Bas van Fraassen and Kit Fine, among many others, now argue that you can *both* deny the law of bivalence, thereby doing justice to the about intuitions the present king of France being bald, *and* hold onto excluded middle, thereby placating the Classical Logic freaks.

Moreover, it's suggested that the mechanisms that allow you to do this make clear how sentences that lack truth values can, as it were, be meaningful in virtue of their potential for having them, thereby solving the first problem as well as the second.

All these glad tidings rely essentially on the method of supervaluations; with which, however, we find that we have grown disenchanted. In this paper, we'll argue that there is something fundamentally wrong with using supervaluation techniques either for preserving classical logic or for providing a semantics for linguistic expressions ordinarily thought to produce truth-value gaps. Our discussion will focus mostly on showing that supervaluation doesn't solve the problem of gapvagueness. But in 'reply 7' we'll make the corresponding case for truth value gaps that arise from presupposition failures. Right from the start, however, we want to emphasize that the objections we're raising are 'philosophical' rather than 'logical'. We have no argument with supervaluations considered as a piece of formal mathematics. Our point will be that if you are partial to the formalism because you're worried about the two philosophical problems we started with, you are in for being badly disappointed.

### *Supervaluations*

So, what is a supervaluation and how exactly is it supposed to preserve classical logic and provide a semantics for vague language? We'll focus largely on Fine's treatment [1975] (however, Lewis [1969], Kamp [1975, 1981], and Dummett [1975] each developed supervaluation-like solutions for (some of) the problems vague language induces). Our exposition of the theory will be as brief and as untechnical as we can manage.

*Classical Valuations:* Suppose we take some first order language  $L$  with monadic vague predicates. Syntax as usual. And let's imagine a (partial) model  $M$  for  $L$  with assignments of extensions to predicates in such a way that some objects in the domain are *definitely in* the extension of a predicate; others are *definitely not in* its extension; and for the remaining objects it is indeterminate whether they are or aren't in the extension. Semantic values for sentences will be built up in familiar ways; some sentences will be definitely true, some definitely false and others without truth-value.

A *classical valuation* closes the truth-value gaps in a partly arbitrary fashion [van Fraassen 1966b, 1968]. On the one hand, an admissible classical valuation  $M^*$  must preserve definite truth and definite falsehood. That is:

If a sentence  $S$  is definitely true in the partial model  $M$ , then  $S$  is definitely true in every admissible classical valuation  $M^*$  derived from  $M$ ; and if a sentence  $S$  is definitely false in  $M$ ,  $S$  is definitely false in every admissible classical valuation  $M^*$  derived from  $M$  [van Fraassen 1966b, p.486].

On the other hand, if [a is F] is neither definitely true nor definitely false in M (because *a* neither definitely is nor definitely isn't in the extension of F in M) then the denotation of *a* must be quasi-arbitrarily assigned either to the extension or anti-extension of F (but not both) in every classical valuation  $M^*$  of M -- this is what Fine calls a *precisification* of the extension of F. In effect, it guarantees bivalence in all acceptable classical valuations of M (thus their name; van Fraassen [1966b]).<sup>3</sup> Certain further coherence requirements ---what Fine calls *penumbral connections*--- are also imposed on these quasi-arbitrary assignments of truth values to sentences in classical evaluations. For example, consider the two vague predicates, 'red' and 'pink'. The umbra of 'red' are those things that are definitely red, i.e., of which 'x is red' is definitely true, while the penumbra of 'red' are those things that are not definitely red, but also not definitely not red, i.e., those things of which 'x is red' is neither true nor false. Similarly for the umbra and penumbra of 'pink'. Notice that the umbras of 'red' and 'pink' exclude each other; their intersection is empty. But 'red' and 'pink' share a penumbral region assuming there are things that are borderline between pink and red.

Precisification achieves bi-valence; precisified predicates only have umbras. However, if the precisification of a predicate is to be meaning preserving, various non-logical (in effect, "conceptual") relations of exclusion and inclusion must be respected. In the present case, since nothing can be both red and pink, a precisification is acceptable only if the new extensions of 'red' and 'pink' do not overlap. This is what Fine calls an "external" penumbral connection [Fine, 1975, p. 276]. He also speaks of "internal" connections [Fine, 1975, p. 276]: If, in a precisification, Al goes into the extension of 'bald' and Bill has the same number of hairs as Al or fewer, then Bill must go into the extension of 'bald' too; and so forth.

We can now say what a supervaluation is. A supervaluation on M combines the results of all the classical valuations  $M^*$  on M; in effect it tells us what all the acceptable classical valuations on M have in common. Thus, a sentence is *supertrue* on a base model M if it is true in every classical valuation  $M^*$  of M; and it is *superfalse* on a base model M if it is false in every classical valuation  $M^*$  of M; and it is neither supertrue nor superfalse otherwise.

So, whatever is true in the original model M will be supertrue in any supervaluation on M, *but there may be supertruths that are not true in the original model*. For example, even if P is neither definitely true nor definitely false in M, any sentence of the form [P or -P] will be supertrue in any supervaluation on M. This is because P will be assigned either true or false in every classical valuation  $M^*$  of M; and the semantics for 'or' and 'not' is standard. In short: even if there are truth value gaps in M, excluded middle is preserved by supervaluation because classical valuations impose bivalence. Similarly, *mutatis mutandis*, for other (propositional) classical logical truths and falsehoods.

Let's now consider how this apparatus applies to the case of vagueness in particular.

## *Vagueness*

As we've seen, vague language is supposed to create special problems for semantics. On the one hand, if there are gapvague sentences bivalence is in threat since gapvague sentences are neither true nor false. And vagueness is also supposed to threaten excluded middle when P is a gapvague constituent of a sentence of the form  $\neg P \vee P$ . Supervaluation apparently copes with both these challenges.

If we say a sentence is classically valid iff it is true in every classical valuation on every model M; and that it is super-valid iff it is supertrue on all models M, then it is provable that a sentence is classically valid iff it is super-valid. Hence, the logic for sentences containing vague predicates can be classical *whether or not* some elementary sentences are gapvague. So, supervaluation "leads to a classical logic for vague sentences" [Fine, 1975, p.265].<sup>4</sup> "The theory is a partial vindication of the classical position. For the truth conditions are, if not classical, then classical at one remove. There is but one rule linking truth to classical truth, viz., that true is true in each of a set of interpretations. This rule is of general application and not dependent upon the nature of language or interpretation. The actual work is done by the clauses for truth in a single interpretation, and those are classical" [Fine, 1975, p. 278].

So much, then, for our exposition. Supervaluations are alleged to provide a semantics for vague language (and, mutatis mutandis, language with non-denoting singular terms), to preserve classical logic in face of truth-value gaps, (and, though we won't discuss this, to solve the sorites (and Liar) paradoxes as well.) It's almost too good to be true. *More* than almost, we're inclined to think.

There are, as the reader is most likely aware, all sorts of criticisms of supervaluation techniques in the literature on philosophical logic. Some doubt that they really succeed in preserving classical logical truths [e.g., von Kutschera 1975]; some doubt that they can provide a solution to the sorites paradoxes (as, presumably, a theory of vagueness ought to do) [e.g., Dummett 1975, pp.310-11; Sanford 1976, p.206; Rolf 1984; Wright 1987; Sainsbury 1988; Williamson 1994]; some doubt that they adequately acknowledge that vagueness is essential to the functioning of natural languages [e.g., Dummett 1975 and Wright, 1987, fn 29]. We will remain (relatively) neutral about all such criticisms; partly because we don't understand a lot of them, but mostly because, as far as we can tell, none of them challenges the *intelligibility* of the idea of a classical valuation of sentences that have truth-value gaps. That's what we propose to do.

*What's seriously wrong with supervaluational accounts of vague languages.*

Our problem with supervaluation it appears to flout what we take to be a platitude; viz., (P):

(P) Conceptual truths must be respected in all classical models, including classical valuations.

We doubt that anybody is likely to be seriously opposed to (P), so we won't propose an extended argument in its favor.<sup>5</sup> Suffice it that honoring (P) is the very motivation that leads supervaluation theorists to postulate "penumbral connections". According to Fine, (and Dummett and Kamp), you need penumbral connections to assure that (eg.) there are no red things in the extension of 'pink' in any classical valuation of English. Surely this must be because all instances of the schema 'if x is red, then x is not pink' are conceptual truths.

So we assume (P). Having done so, we propose to argue that there are conceptual truths about gapvague sentences which:

- i. P demands that Classical valuations respect and,
- ii. if respected, preclude certain of the precisifications that Fine, Dummett, Kamp et al say are needed if a supervaluation is to recover truths of classical logic.

We take this to imply that there is no way of assigning supervaluations to gapvague sentences that respects *all* the necessary truths, *both logical and conceptual*. So supervaluation can't be the right semantics for gapvague sentences.

Assume, what we suppose is plausible, that there is no fact of the matter about whether a man one ninth of whose head is covered with (his) hair is bald. (If you don't like this fraction, feel free to pick a different one; all that we require is that sentences of the form 'if x has (fraction) of his head covered with hair, then x is bald' are definitely gapvague.) We claim that if there is no matter of fact about whether someone one ninth of whose head is covered with hair is bald, then it is *non-contingently* the case that there is no matter of fact about whether someone one ninth of whose head is covered with hair is bald: If it's unsettled in the actual world then it's unsettled in every world. If you doubt this, ask yourself what facts about the world (or about English, for that matter) would convince you that you were wrong and that, by gum, people one ninth of whose heads are covered with hair are definitely bald after all. Remember that baldness doesn't have a 'hidden essence'. You couldn't discover empirically that hairless people aren't bald in the way you could (maybe) discover empirically that XYZ isn't water.

Now consider Al. If it's non-contingent that there is no fact of the matter about whether someone one ninth of whose head is covered with hair is bald, then every model in which Al has one ninth of his head covered with hair is ipso facto a model in which it is indeterminate whether Al is bald. That is, the sentence 'Al is bald' (hereinafter "The Al Sentence") is indeterminate in every such model. But the point of precisification it to impose bi-valence; it intends to *exclude* models in which the Al sentence is neither

definitely true nor definitely false. To do so, it requires that we consider indefinitely many models in which Al has hair on one ninth of his head and the Al sentence is true, and that we consider indefinitely many models in which the Al has hair on one ninth of his head and the Al sentence is false. But now we see that there can be no models of either kind: On the one hand, the connection between having hair on one ninth of one's head and being neither determinately bald nor determinately not bald is conceptual, and on the other hand, (P) tells us that acceptable models must respect all the conceptual truths.<sup>6</sup>

So you can't precisify a language which contains the sentence (not, of course, the mere form of words; a form of words can mean what you will) 'Al is bald'. So you can't precisify English. So you can't supervalue the gapvague sentences of English *and the reason that you can't is precisely that they are gapvague*. Gapvagueness is, as one might say, an enduring trait of the sentences that have it. You can't both think it away and preserve the semantics of the sentences that you're thinking about.

### *Replies and replies to replies*

We think this argument undermines the very idea of precisifying vague sentences, so we could stop here. There are, however, some interesting lines of reply that we've gleaned from the literature or heard in conversation. We find them very unconvincing; saying why we do should help clarify our view of the conceptual geography. The rest of the paper is devoted to that.

*Reply 1:* Someone might complain that our worry has nothing to do with supervaluation theory per se; flouting conceptual truths is something that's done all the time in classical model theory.

Suppose we ask a student to evaluate the argument: [Something is a round square; so, everything is a round square.] We expect the verdict 'formally invalid in predicate logic.' Are we thereby requiring the student to do what can't be done? Of course not! But if we can reasonably demand the consideration of models that flout the conceptual truth that nothing square is round, why can't we equally reasonably demand consideration of models in which a necessarily vague sentence has a truth value?

But this goes much too fast. To be sure, the argument we've given the student to evaluate contains a premise that's conceptually false. But *that it does isn't what makes the argument invalid in predicate logic*. What makes it invalid is that it wants to infer from 'some x is F' to 'every x is F'. In effect, when we ask the student to evaluate the argument, we're asking him to attend only to those of its features to which predicate logic is designed to be sensitive: relations among the connectives and among the quantifiers of the sentences; relations of identity and difference among their predicate structures (e.g., whether the premise and conclusion have the *same* predicate -- *whatever that predicate may mean*), and so forth. Since, by stipulation, the argument's validity turns *only* on these

features, the occurrence of 'round square' in the premise and conclusion is *inessential*. The assessment would have gone through in *exactly the same way* if the argument had been about red circles, or old houses, or prime numbers. When we consider the formal validity of an argument containing 'round square', we don't *flout* the meaning of the expression; we just *ignore* it.

So, then, to repeat the question that the objection started with: If predicate logic can abstract from the conceptual incoherence of talking about squares being round, why can't supervaluation abstract from the conceptual incoherence of talking about the AI sentence being definitely true? Answer: The vagueness of the AI sentence depends on its *nonlogical vocabulary*; so a theory of vagueness can't abstract from the semantics of the nonlogical vocabulary any more than predicate logic can abstract from the semantics of 'and' or 'some'. If what you care about is the "truth conditions for a vague language" [Fine, 1975, p.265] you can't ignore the lexical features *in virtue of which* language is vague. It's the semantics of the nonlogical vocabulary that a theory of vagueness is supposed to be a theory *of*.

Look, if it is legitimate to save excluded middle by flouting some of the conceptual truths that arise from the nonlogical vocabulary, why not call a spade a spade and say that what makes a sentence logically true has *nothing at all* to do with the semantics of its nonlogical vocabulary; that logical truth is determined entirely by logical form. This is what van Fraassen calls the "naive" view that "questions of validity can be decided on the basis of syntactic form alone" [van Fraassen 1966b, p.483]. Well, if it's pretty clever and sophisticated to ignore the imprecisifiability of a necessarily vague term, why wouldn't it be *really* clever and sophisticated to ignore the rest of the semantics of the non-logical terms as well, and hold that 'either *you are bald if you have one ninth hair cover* or *not (you are bald if you have one ninth hair cover)*' is true because it's of the form  $P \vee \neg P$ ?

*Reply 2:* 'You say that if it is indeterminate whether 1/9 hair cover is bald, then it is necessary that it is indeterminate whether 1/9 hair cover is bald. But that can't be right. Notice that it is perfectly coherent for someone to stipulate (perhaps in a certain context, or for a purpose) that 1/9 hair cover is bald. On your account that couldn't be so, since the proposition thus stipulated would be necessarily false.'

We think this shows what we are anyhow inclined to believe: Pace Tappend (1993), it's unwise to take stipulation as a model for precisification. In particular, you can't assume that if it's alright to stipulate that P, then there must be an acceptable model in which P is true.

The reason you can't is that, whereas it's common ground that precisification is required to honor the penumbral conditions, acceptable stipulations often flout them. Often enough, when you stipulate that a is F, what's being mandated is really just that *a shall be counted with the Fs* (in the context or for the purpose at hand.) When that's its force, stipulation

can defy patent truths; including patent conceptual truths. 'Married men separated from their wives shall count as bachelors [indeed, *are* bachelors] for the purposes of the census;' 'Ketchup counts as a vegetable for the purposes of the Republicans' school lunch program;' 'Coaches count as players for the purposes of the official roster;' 'Neither Jews nor Italians count as minorities for the purposes of affirmative action;' and so forth.

We have no positive account to offer of what makes, or can make, a stipulation acceptable. We'd be surprised if an account of any substance could be of any generality. Suffice it that stipulations needn't, and often don't, preserve the semantic features on which the identity of an expression depends. That you can stipulate that 1/9th covered is bald does not, therefore, imply that there is a model in which that is so.

*Reply 3:* Here's a positive proposal about what it is for a language to endure in the face of semantic change: definite extensions (i.e., the extensions in the umbra) have to be preserved.

[L]anguage can retain its identity upon precisification... The identity of language is visible, as it were, in the permanence of recorded truth...[I]et the actual meaning of a simple predicate, say, be what helps determine its instances and counterinstances. Let its potential meaning consist of the possibilities for making it more precise. The the point is that the meaning of an expression is a product of both its actual and potential meaning. In understanding a language, one has thereby understood how it can be made more precise [Fine, 1975, p.277].

The idea is that supervaluation answers the question 'How can classical logic be preserved for a vague language' *because* it answers the question 'how can a sentence that can't be semantically evaluated mean anything at all?' What preserves the meaning of a sentence under precisification *just is* that precisification preserves definite truth and falsity. Construed as an answer to our sort of objection, this reply simply refuses to niggle about the conceptual probity of assigning truth values to gapvague sentences. Rather, it proposes a perfectly general *sufficient* condition for distinguishing precisification from stipulation (precisification preserves extensions in the umbra; stipulation doesn't), and it defends the condition by arguing that it draws the line at an intuitively acceptable place. In effect, this reply says that flouting P in the precisification of gapvague terms is ok because the gapvagueness of a term isn't essential to its identity:

...a sentence does not become indefinite upon being made more precise. This is, perhaps, partly definitional of 'making more precise'. For what distinguishes this operation from a mere change in meaning is that it preserves truth value. [Fine, 1975, p.275].

Here's Dummett in a similar vein:

For every vague statement, there is a certain range of acceptable ways of making it definite, that is, of associating determinate truth-conditions with it. A method of making a vague statement definite is acceptable so long as it renders the statement true in every case in which before, it was definitely true, and false in every case in which, before, it was definitely false [Dummett, 1975, pp.310-11].

The trouble with this line, according to us, is that the distinction between stipulation and precisification that it proposes isn't plausible after all. We would have thought that what makes the precisification of a predicate (or, *mutatis mutandis*, any linguistic change) meaning preserving is *not* that it conserves unaltered the (determinate) truth values of sentences that contain the predicate, but rather that it conserves the necessary truths that those sentences express, *conceptual necessities included*.

Here's a thought experiment that illustrates how these two criteria for meaning-preservation can diverge. Suppose that, in aid of calendar reform, we decide to skip a day, so the third of next month is a Tuesday instead of a Monday. The reform is thus *unconservative* about the determinate extension of 'Tuesday', but the intuition (anyhow, our intuition) is that it *is* conservative about what 'Tuesday' means. Now imagine a reform that institutes a six day week, with Wednesdays following Mondays. This violates (inter alia) *Weeks have 7 days* and *Wednesday is the third day after Sunday*, both of which are presumably conceptually necessary. In this case, the (our) intuition is that the meanings of the terms in this part of the language haven't, to use Fine's term, "endured the change."

If you don't like the example, so be it: For, the point it is supposed to illustrate isn't really moot in the current discussion. As we remarked above, Fine et al put the "penumbral connections" in place precisely to insure that conceptual necessities like *if red, then not pink* are preserved under precisification. That must be because they think precisification ought to preserve something *more* than determinate extension; specifically, it must be because they think that precisification ought to preserve conceptual necessities. But that it is indeterminate whether someone who has one ninth hair cover is bald *is* conceptually necessary, and so must be preserved under any precisification that illuminates the meaning of 'bald'. Which, however, it isn't if there are acceptable models in which AI has one ninth hair cover and the AI sentence is true. Such a model fails to meet the conditions for "enduring" language, in just the same way, and to just the same extent, as a model in which 'This red thing is pink' is true. *Respecting determinate extensions is not, after all, a sufficient condition for meaning to be preserved under linguistic change*. This should surprise no one who thinks that meaning and conceptual necessity must somehow be intrinsically related.

*Reply 4*: "Alright, alright, so strictly speaking, 'bald' doesn't mean *bald* in a language in which 'is bald' is determinately true of someone with one ninth hair cover. So the language

for which precisification preserves the truths of Classical Logic isn't, strictly speaking, English. But, gee, it's *a lot like* English. Like enough so that we can learn interesting things about what 'bald' means in English by attending to the behavior of its not-quite-English counterpart."

Reply 3 said that the necessities that precisification violates aren't constitutive of the meanings of gapvague terms. Reply 4 admits that they are but is prepared to live with the resulting equivocations. The literature doesn't always respect the difference between Reply 3 and Reply 4, but sometimes it does. For example:

[U]nder the pressure of their own use, the meanings of terms will need to change. The terms, in their old sense, will not be adequate to express the new truths, pose the next questions, make the right distinctions. Now clearly it is convenient that the changes in meaning be conservative, that the true records before the change remain true after the change [Fine, 1975, p.275].

Or, consider Sainsbury:

'Heap' is a vague predicate but it could be replaced by a sharp predicate.... Indeed such replacement is the only way forward if someone relentlessly pursues the question, asked of a penumbral object, whether it is a heap or not. We have to say: There is no answer to that question, but we could introduce new predicates, rather like 'heap,' of which we could answer that question [Sainsbury, 1988, pp.38-39, our emphasis];

or:

[We] 'sharpen' the predicate 'heap' by using the new predicate 'newheap'...We will say that [a sentence containing a vague predicate] is (definitely) true iff it is true however its vague predicates are sharpened [Sainsbury, 1988, p.33].

Oh, well, one chap's light is another chap's darkness. It is unclear to us how the vagueness of English expressions is to be illuminated by investigating the homophonic expressions in a language which is *not* English and *none* of whose terms is vague. If it's clear to you, you're welcome to it.

*Reply 5:* "A model is acceptable iff every sentence that is determinately true is true in the model; what's the big sweat?" (Our colleague Vann McGee once said something like this to us, only more politely.) Answer: This criterion of acceptability presupposes a viable notion of *same sentence*; it's got to be the *same* sentence that is both determinately true and also true in the acceptable model. But sentences are *semantic* entities, so a necessary condition for their identity is (presumably) the preservation of the conceptual truths that they enter into. If acceptable models can flout conceptual truths, the proposed criterion is *not* satisfied by precisifications. (And if, per contra, sentences *aren't* semantic entities, it

beats us how the fact that a sentence is true in a model iff a *homophonic* sentence is determinately true could have anything to do with the acceptability of the model.)

*Reply 6:* "You guys think that precisifying extensions would result in (e.g.) the sentence 'Anyone with one ninth hair cover is bald' being assigned true (i.e., definitely true) in some classical evaluations; but that's unsubtle. What's really assigned is something like true-in-a-classical-valuation; or better, t-in-a-classical-valuation. Who says a sentence can't be both necessarily indeterminate and t-in-a-classical-valuation? 't-in-a-classical-valuation' is *my* word, so *I* get to decide what it entails."

Fine! But if t-in-a-classical-valuation is completely independent of truth *tout court*, i.e. of so-called definite truth, then how can the fact that 'Al is bald' is t-in-some-classical-valuations and f-in-others make the complex sentence 'Either Al is bald or Al is not bald' definitely true? We thought the claim was that the disjunction is definitely true because at least one of its disjuncts is true in every classical-valuation. Why on earth should having one disjunct that is 't-in-a-classical-valuation' in every valuation make a disjunction definitely true?

To put the same point slightly differently, the present objection is that 'true' in 'true-in-a-classical-valuation' is an orthographic accident, not to be confused with the 'true' in 'definitely true.' But then, how come it is possible to define 'definitely true' in terms of it? Having a friend who is dogmatic is quite unlike having a friend who barks, so it would be a surprise if you could construe the one in terms of the other. Likewise, according to the present objection, being 't-in-a-classical-valuation' is quite unlike being 'true' since a sentence that is necessarily indeterminate can be 't-in-a-classical-valuation'. But (surprise, surprise!) being t-in-*every*-classical-valuation entails being definitely true all the same. How are we to understand this? And if it is not to be understood but merely to be admired, *in what sense has supervalutation shown us how vague language works?*

It has been suggested to us that we should think of 't-in-a-classical-valuation' as being defined not explicitly but in use (a la Wittgenstein rather than Russell, we suppose). If we meditate on how the predicate functions across an exhaustive set of classical valuations, we'll come to understand it; and the logic and truth conditions of vague language will thereby be made clear to us. Well, maybe. But, so far, what strikes us when we meditate on the behavior of 't-in-a-classical-valuation' is that, although we're told that it doesn't *mean* true, and although we take the hyphens to say that it doesn't *contain* 'true', still, atomic sentences that are t-in-classical-valuations contribute *exactly what true sentences would* to the semantics of the complex sentences that contain them. Could this be an accident?

Really, of course, for a sentence to be true in a model is for it to be *true* in that model; it's for the sentence to bear the same relation to the model that a (definitely) true sentence bears to the actual world. The intuitive force of supervalutation depends on this in all sorts

of ways. For example: if you break the connection between 't-in-a-model' and 'true in a model', there's nothing left to guaranty that whatever is t-in-a-model is ipso facto *truth evaluable*, either in that model or, indeed, *at all*. So, what's to stop 'the number two is feeling blue' or 'colorless green ideas sleep furiously' getting t-evaluated along with 'the cat is on the mat,' 'the present king of France is male' and 'Al is bald'?

In fact, once you break the connection between 't-in-a- model' and 'true in a model,' what is there to make being-t-in-a-model a *semantic* property?<sup>7</sup> And, if there is *nothing* to make it a semantic property, then presumably things that have *no* semantic properties (tables and chairs, as it might be) could be t-in-a-model too. Isn't there pretty clearly something wrong with a theory of vagueness that is compatible with tables and chairs being vague?<sup>8</sup> We'd supposed that supervaluation was going to explain how some things that can't be semantically evaluated (eg. gapvague sentences) can nonetheless be meaningful in ways that other things that can't be semantically evaluated (eg. tables and chairs and '2 is blue') are not. We'd understood that was one of the things that the ads commend the theory for. But apparently not.

*Reply 7: Outside the base model, elementary sentences are holophrastic.* (This one will take a little longer.)

Supervaluations were first employed by van Fraassen to provide a semantics for free logic, i.e. logic that allows for genuinely non-denoting singular terms [van Fraassen 1966a, 1966b; van Fraassen and Lambert 1967]. Subsequently, van Fraassen applied the notion of supervaluations, or a generalization therefrom, to analyze the relation of presupposition and the semantics of sentences that lack truth-value on account of presupposition failures. (He then applied the conceptual apparatus to such semantic paradoxes as the Liar and the Strengthened Liar [van Fraassen 1968, 1969, 1970].) Since the objection we're about to consider builds on van Fraassen's account of supervaluation, and since his discussion centers not on vagueness but on truth-value gaps that occur when sentences contain vacuous (non-denoting) singular terms, we now turn to these. Whereas the supervaluation treatment of vagueness is widely agreed to be moot, its use in respect of problems about non-denoting singular terms has become something of a received practice.

As van Fraassen conceives of a classical valuation, if the name 'Pegasus' or the definite description 'the present king of France' does not denote in a partial model *M*, there are *no* constraints on what truth-values may be assigned to atomic sentences containing them in a classical valuation *M\** on *M*.(fn.). That is, truth-values are assigned entirely arbitrarily to sentences like 'Pegasus has wings' or 'The present King of France is male.'<sup>9</sup> Both are assigned true in some classical valuations *M\** on *M*, false in others, and the supervaluation story about complex sentences like 'Either the present king of France is male or the present king of France isn't male' goes through as before.

But now our sort of worry would appear to arise again. Just as some predicates are

necessarily vague, some singular terms necessarily fail to denote; e.g., 'the man who is taller than himself' (and also 'Pegasus' if Kripke's story about names is right [1980]). So it looks like van Fraassen classical evaluations are in the business of assigning (nonce) truth-values to sentences whose putative referring expressions *can't* refer; hence, to sentences which *can't have* truth-values. How are such assignments to be understood?

(What we take to be) van Fraassen's answer to this constitutes Reply 6: What truth-value a sentence has in a given  $M^*$  is independent of any semantic contents of either the predicates or the singular terms that it contains. In effect, sentences like 'The present king of France is male' and 'Pegasus has wings' and 'The man who is taller than himself is Russian' are treated as mere surds; i.e., as having assigned truth-values but no internal semantic structure. So, for example, on van Fraassen's account, it is quite consistent for both 'The man who is taller than himself is Russian' and 'Necessarily there is no man who is taller than himself' to be true in a classical valuation  $M^*$  (indeed, the second must be true in *every* classical valuation  $M^*$  since it is definitely true and classical valuation preserves definite truth/falsity.) van Fraassen wants us to take seriously the idea that no semantic value is assigned to the non-denoting singular terms in his classical valuations.

Well, of course, you can *say* that if you like; but, if you do we promise not to understand you. First blush, a sentence can't have a truth-value if (i.e., in a model in which) its syntactically singular referring expressions don't refer. That's *why* people think that 'the present king of France is tall' lacks a truth value in the actual world. But then, if a sentence contains a singular term that fails to refer *necessarily*, it can't have a truth value in any model. A fortiori, it can't have an arbitrary truth-value in any model. So it can't be supervalued.

Possible reply: "No, no, you miss the point when you speak of 'syntactically singular referring expressions.' I'm saying that 'The man who is taller than himself is Russian' has no internal structure in  $M^*$ , *semantic or syntactic*. Or, rather, since it has no internal syntactic structure in  $M^*$ , a fortiori it has no internal semantic structure in  $M^*$ . As far as the assignment of a truth value in  $M^*$  is concerned, the sentence is holophrastic. Its truth value doesn't depend on its parts because it hasn't any parts for its truth-value to depend on."

We can hardly believe our ears. By what criterion is it *the very same sentence* that does contain 'the man who is taller than himself' as a syntactic constituent when it is evaluated in  $M$  but does not contain 'the man who is taller than himself' when it is evaluated in the precisification  $M^*$ ? Are you saying that what syntax a sentence has depends on what model it is evaluated in? Geez! Say it isn't true.

The point we're making about objection 7 is really continuous with the one that we made about the supervaluation treatment of gapvague sentences; both propose to sacrifice a kind of semantic innocence that it surely would be well to retain. In the former case, if the

assignment of truth values to a gapvague sentence is allowed to violate conceptual necessities, how can it be said that the meaning of the sentence is preserved under precisification? Analogously: If a sentence with a singular term gets its truth value in one way (namely, in virtue of its compositional structure) in a model where the term denotes, but in quite a different way (viz arbitrarily) in a model where the term doesn't denote, how can it be said that the meaning of the sentence is preserved from one model to the other? What becomes of the innocent assumption that if God were to provide us with a present king of France, He wouldn't thereby change what `the present King of France' *means*?

van Fraassen says that "a classical valuation gives a faithful picture of which sentences are true and which false in some (possible) situation" [1969, p.71, our emphasis]. Faithful to what, we wonder? Not to truth values since they are assigned arbitrarily; not to internal semantic structure, since it is ignored. Not to internal syntactic structure, since it is ignored too. Not to truth conditions, since, *prima facie*, some of the sentences we're concerned with ('The man who is taller than himself is Russian') are necessarily indeterminate, so they don't have truth conditions. All that the analysis is *really* faithful to, as far as we can tell, is the logical form of the complex sentences in which non-denoting expressions occur. See above, Reply 1 and Reply 3, for the corresponding point about vagueness.

No English sentence of the form *The man who is taller than himself is X* can be true. This is conceptually necessary, so it is true in every model in which the semantics of English singular terms is respected. To insist that there is a model in which some *other* sentence --- one which does *not* contain the expression of `the man who is taller than himself'--- is true, is to tell us nothing about the semantics of singular terms in English. What it tells us is just that you have decided to change the subject.

*Conclusion:*

The intuitive appeal of the supervaluation technique comes, we think, from cases like `The present king of France is a male.' Since it's only contingent that there is currently no king of France, it's only contingent that this sentence lacks a truth value. Where a truth-value gap is merely contingent, we know pretty well what it would be like to fix the world so as to make the gap go away. The situation is less clear in the case of sentences with vague terms, because it's less clear that it's the world's fault that they lack truth-values; maybe you can't make them precise without changing just the thing about them that you're trying to study.

That certainly seems to be so when the lack of truth-value is conceptually necessary; conceptual necessity is itself presumably a semantic phenomenon, at least *inter alia*. Which is to say that it must be preserved by any operation that claims to be meaning-preserving. For a semantic theory to fool around with the *constitutive* properties of an expression is for it to describe the semantics of some *other* expression. To solve a semantic problem by

pretending away necessary truths is only to pretend to solve the problem.

## APPENDIX

### Bencivenga's Nonholophrastic Treatment of Empty Descriptions.

Like us, Ermano Bencivenga has misgivings about van Frassen's not making the evaluation of elementary sentences depend on their internal structure. Bencivenga remarks, for example, that the identity symbol is as good a logical constant as the disjunction symbol, but the former will be invisible if the internal structure of elementary sentences is ignored.

We entirely approve of Bencivenga's attempt to avoid holophrasism, but his theory cannot cope with our chief worry, which is how to avoid valuations that assign referents to necessarily nonreferring expressions *while also avoiding valuations that change the subject*. A full treatment would need more space than we've got, so we'll scant the details.<sup>10</sup>

Bencivenga's assignment of denotations to definite descriptions is perfectly orthodox: In any model, 'The-F' will have for its denotation the unique member of the domain that satisfies F, if there is one. If there is no such unique member in a model, then a sentence containing 'The-F' will lack a truth value in that model. Since complex sentences are also evaluated in the usual way, the truth value even of cases of the excluded middle, like 'The winged horse is white or the winged horse is not white', will be undefined in *some* models. This means that Bencivenga can't hold that a sentence containing 'The-F' is supertrue just in case it's true in every extension of the base model M. Instead, he will restrict the quantification over models *just to those extensions in which a value for the sentence is defined*. This spares him the need to assign arbitrary denotations to empty descriptions.

In effect, what Bencivenga does is this: As usual, let a model M be a pair (D, f), where D is the domain and f is a function that assigns extension to terms in the familiar ways. If  $M = (D, f)$  and  $M^* = \langle D^*, f^* \rangle$ , then  $M^*$  is an *extension* of M just in case D is a subset of  $D^*$  and f is a subset of  $f^*$ . Since there are no winged horses in any customary model M for English, the intersection of  $f(\text{'winged'})$  and  $f(\text{'horse'})$  will be empty in all the customary models and the denotation of 'the winged horse' will accordingly be undefined in all the customary models.

However, it is easy to imagine a model  $M^*$  that is an extension of M in which:  $f^*(\text{'winged'}) = f(\text{'winged'})$  and  $f^*(\text{'horse'}) = (f(\text{'horse'}) + 1 \text{ white seagull})$ . For such an  $M^*$ ,  $f^*(\text{'the winged horse'}) = \text{that seagull}$ , and 'the winged horse is white' is true in the model.

This model is ipso facto relevant to the supervaluation of 'the winged horse is white'. So too, of course, is a model in which  $f(\text{'horse'})$  is extended to include a black crow rather than a white seagull and 'the winged horse is white' is false.

So there are some extensions of the customary model in which the value of 'the winged horse is white' is defined and the sentence is true and there are some extensions in which its value is defined and it is false. It follows that this sentence is neither supertrue nor superfalse in the customary model. But 'the winged horse is white or the winged horse is not white' *will* be supertrue in the customary model since it is true in every extension where the value of 'the winged horse is white' is defined.<sup>11</sup>

The reader will, no doubt, have anticipated our objection: If all that's required of  $f^*(\text{'horse'})$  is that it include  $f(\text{'horse'})$  as a subset, then extensions won't, in general, be meaning preserving. Patently, they won't be if  $f^*(\text{'horse'})$  includes elements to which it is *conceptually necessary* that 'horse' does not apply. Consider an extension in which 'the winged horse is white' is true because  $f^*(\text{'horse'})$  contains a white seagull. There would seem to be only two possibilities: either it's a model in which some seagulls are horses, or it's a model in which 'horse' doesn't mean *horse*. Since we're inclined to think that it's some sort of necessary truth that seagulls aren't horses; we think that the evaluation of 'the winged horse is white' as true in this model must involve an equivocation on 'horse'. (If, however, our modal intuitions strike you as wrong, the point can be adjusted to whatever ones suit your fancy. Eg. let the sentence to be evaluated be 'the prime horse is equal to 2' and let  $f^*(\text{'horse'}) = f(\text{'horse'}) + 2$ .)

If 'horse' doesn't mean *horse* in these sorts of models, then 'the winged horse' doesn't mean what it customarily means; that is, *it doesn't mean what it means in the models that represent English*. In that case, it's hard to see how a notion of supertruth that's defined relative to such models could be germane to whether English sentences of the form 'The winged horse is F or the winged horse is not F' are true. This is, of course, the same dilemma that we kept encountering in the body of the text: Where truth value gaps are necessary, the choice supervaluation offers is to ignore salient structure or to equivocate. Neither alternative strikes us as awfully attractive.

## Notes

1. For simplicity, we assume that sentences are the bearers of semantic values, meanings and the like. It's notorious that there are more plausible candidates, but nothing in our argument will turn on this.
2. It is, of course, implausible that *all* sentences with vague predicates lack truth value. 'Bald' is paradigmatically vague, but 'Yul Brenner was bald' is definitely true. Generally, when we speak of vague sentences, we have in mind sentences that have truth-value gaps in consequence of their vagueness. For convenience, we'll call such sentences 'gapvague'.
3. "Thus the classical valuations do not admit truth-value gaps; bivalence holds for them. They proceed on the assumption that there is no possibility of a failure of presupposition" [van Fraassen, 1969, pp.70-71].
4. "[T]he supervaluation technique...was motivated in large part to reflect the intuition that there is difference among such [i.e.,  $a \& a$ ,  $av-a$ ,  $a \& a$ ] cases: if  $a$  lacks a truth-value, so should  $a \& a$ , but  $a \& a$  should nevertheless be counted as definitely false and  $av-a$  as definitely true" [Kamp and Partee, ms].
5. Tappenden (1993) suggests that it's ok if a penumbral sentence isn't true in a model, just as long as it also isn't *false* in that model. Since this weak reading of P is strong enough for our polemical purposes, we don't propose to argue the issue.
6. It has been suggested to us that the conceptual truths that we want precisification to honor are somehow "metalinguistic" or "higher level" in a way that the one's about 'red' and 'pink' and such are not. It's unclear to us why it would matter if this were so; and, anyhow, as far as we can see it isn't. The crucial consideration is the, as it were, object level fact that *you can't make a man more (/less) bald without altering his hair-to-head-ratio*. So if there is *any* valuation on which the AI sentence is true (/false/indeterminate) and AI's head-to-hair ratio is  $m/n$ , then the AI sentence is true (/false/indeterminate) in *every* valuation in which AI's head-to-hair ratio is  $m/n$ .
7. To be sure, having a truth value in a model is a *sufficient* condition for being t-(or-f)-in-that-model. But that's not enough to make being t-(or-f)-in-the-model itself a semantic property. Being butter is sufficient for being yellow; it doesn't follow that yellow is a dairy product.
8. Like classical evaluations of sentences, classical evaluations of tables and chairs preserve definite truth and falsity. As it happens, they do so vacuously.
9. Exclude identity statements; and existence statements, if there is a primitive predicate

for existence.

10. Though his theory has altered slightly over the years, we will restrict our discussion to Bencinvenga's accounts of (1980a,b) and (1991). Bryan Skirms (1968, 1970) has independently advocated nonholophrastic evaluation. We won't discuss Skirms, but we believe our criticisms apply to his proposal too.

11. This is, in fact, a considerable simplification of Bencivenga's treatment, but it conveys the underlying intuition. It will suffice for our purposes since his notion of extension is *less* restrictive than the one that we've just described.

### *Bibliography*

Bencivenga, E., "Truth, Correspondence, and Non-denoting Singular Terms, *Philosophia*, 1980a: 219-229.

Bencivenga, E., "Free Semantics for Definite Descriptions," *Logique et Analyse*, 1980b: 393-405.

Bencivenga, E. "Free Logic," *Handbook of Philosophical Logic*, v. 3, (D. Gabbay & F. Guethner, eds.), 1986.

Dummett, M., "Wang's Paradox," *Synthese* 30, 1975. Also in *Truth & Other Enigmas*, 1978: 248-268.

Bencivenga, E. "Free Semantics" in *Philosophical Applications of Free Logic*, (K. Lambert, ed.) 1991.

Bencivenga, E., Lambert, K. and van Frassen, B., *Logic, Bivalence and Denotation*, Ridgeview Publishing Company, Atascadero, California, 1986.

Skyrms, B. "Supervaluations, Identity, Existence and Individual Concepts," *Journal of Philosophy*, 69, 1968: 477-482".

Skyrms, B. "Return of The Liar: Three-valued Logic and The Concept of Truth." *American Philosophical Quarterly*, 1970.

Fine, K., "Vagueness, Truth and Logic," *Synthese* 30, 1975: 265-300.

Kamp, J.A.W. "Two Theories about Adjectives" in *Formal Semantics of Natural Language*, ed. by E.L. Keenan, 1975:123- 155.

Kamp, H. "The Paradox of the Heap," in *Aspects of Philosophical Logic*, ed by U. Monnich, 1981:225-277.

Kamp, H. and B. Partee, "Prototype Theory and Compositionality," *Cognition*, forthcoming.

Kripke, S., "Outline of a Theory of Truth," *Journal of Philosophy*, 72, 1975:690-716.

Kripke, S., *Naming and Necessity*, Cambridge, MA: Harvard University Press, 1980.  
von Kutschera, Franz, "Partial Interpretations, in *Formal Semantics of Natural Language*, ed. by E. L. Keenan, 1975.

Lambert, K., "Logical Truth and Microphysics" in *The Logical Way of Doings Things*, ed., K. Lambert, Yale University Press, 1969.

Lewis, David, "General Semantics," in *Semantics for Natural Language*, eds., D. Davidson and G. Harman, D. Reidel: Dordrecht, 2nd ed., 1972:169-218.

Rolf, Bertil, "Sorites" "Sorites" *Synthese*, 1984:219-250.

Sainsbury, R. M., *Paradoxes*, Cambridge University Press, Cambridge, 1988.

Sanford, D., "Competing Semantics of Vagueness: Many Values Versus Super-Truth," *Synthese*, 33, 1976:195-210.

Tappend, J. "The Liar and Sorities Paradoxes: Toward A Unified Treatment," *Journal of Philosophy*, XC, 1993: 11, 551-557.

van Fraassen, B.C., "Singular Terms, Truth-value Gaps, and Free Logic," *Journal of Philosophy*, 63, 1966b:481-95.

van Fraassen, B.C. and K. Lambert, "On Free Description Theory," *Zeitschrift fur mathematische Logik und Grundlagen der Mathematik*, 1967.

van Fraassen, B.C., "Presuppositions, Implications, and Self-Reference," *Journal of Philosophy*, 65, 1968:136-52.

van Fraassen, B.C., "Presuppositions, Supervaluations, and Free Logic," in *The Logical Way of Doings Things*, ed., K. Williamson, Timothy, Vagueness, Routledge, London and New York, 1994, Chapter 5, pp. 142-164.

Wright, Crispin, "Further Reflections of the Sorites," *Philosophical Topics*, 1987:227-290.