Cut-offs and their Neighbors

Achille C. Varzi
Department of Philosophy, Columbia University, New York

(Final version published in J. C. Beall (ed.), *Liars and Heaps: New Essays on Paradox*, Oxford: Oxford University Press, 2003, pp. 24–38.)

I agree with Priest: the sorites is a very hard paradox, possibly harder than the liar.[1] For the liar can be isolated, whereas the sorites is everywhere and can take us anywhere. And I agree that the paradox is so hard because it systematically imposes upon us the existence of unbelievable or otherwise unacceptable cut-off points. No solution can avoid explaining why this happens. But I have two points of disagreement with Priest's account of the matter. The first concerns the emphasis that he places on the working of natural language. And the second, more important problem concerns the line of explanation that Priest offers of the reason why we are stuck with cut-off points—hence his tentative solution to the paradox.

1. Sorites Progressions and Linguistic Vagueness

Let's focus on the forced march sorites, in the form given by Priest. We have got a series of questions $q_0 \dots q_n$, each of the form 'Is it the case that β_i?', and the facts are such that the answer to the first question, a_0, is different from the answer to the last question, a_n. Given these facts, there is no way out: the series of answers must involve a cut-off point. No matter what form the answers take, and no matter how we feel about the relative indistinguishability of any two successive β_i's, if $a_0 \neq a_n$ there must be a first $k > 0$ such that $a_0 \neq a_k$.

Now, Priest takes the problems into which this scenario may lead us to be problems posed by the vagueness of our language. So where exactly does language play a role in this scenario? Not in the answers, for these can be anything we like, including straight Yes or No answers. Besides, as Williamson made clear a while ago, the answers are actually irrelevant since the problem concerns what

[1] Unless otherwise specified, all references are to Priest (2003).

one should believe, not just what one should say.[2] So language must play a role in the questions, the q_i's, hence in whatever bits of language we use to express the statements corresponding to the β_i's. (The locution 'Is it the case that...?' is ultimately redundant.) In Priest's example, each β_i is a statement to the effect that you are b_i, namely, an object that comes from you, b_0, by i successive replacements of one molecule of your body with a molecule of scrambled egg. In other words, each β_i is an identity statement of the form

(1) $b_0 = b_i$.

The notion of replacement, we may suppose, is defined with precision, for the predicate 'scrambled egg' that occurs in our description of the procedure can be defined in precise terms. And we may also suppose that your body, b_0, is defined with similar precision, though in actual circumstances this may be a hard thing to do. Accordingly, we can safely assume that in each case the two terms flanking the identity predicate in (1) have a precise semantics. That leaves only the identity predicate itself; no other bit of language is involved in our questions. But don't we know exactly what this predicate stands for? It stands for the identity relation. So either we say that there really are many, slightly distinct candidates for this relation, at least as we conceive of it when we use the identity predicate in ordinary discourse, or else language plays no role at all in the picture and the burden of the sorites would be entirely on the identity relation itself. (We would then have a case of ontological vagueness.) Thus, when Priest says: "How can a single molecule make a difference? … [V]ague predicates just don't seem to work like this", he is making a claim that goes beyond the data. Vague predicates play no role in this scenario except for identity. At most we can say: "The identity predicate doesn't work like this". And we can say that only if we reject the thought that the burden of the sorites is on the identity relation.

Now, I am happy to reject that thought, so let's put ontological vagueness on a side. The question is, what are we to make of the fact that the burden of the forced march sorites is entirely on the identity predicate? One obvious response is that this is just a byproduct of the particular example under discussion. After all, in the scenario described by Priest the questions involve identity statements, but they could be questions of a different sort. For example, the q_i's could be questions of the form 'Is it the case that β_i?', where each β_i is a statement of the form

(2) b_i is a heap,

[2] Williamson (1994: 14).

with b_i defined in the obvious way. In that case the reasoning above would show that the burden of the sorites is on the predicate 'heap', not the identity predicate. So by generalization we would get to the desired point: generally speaking, the sorites forces us to recognize the existence of cut-offs in the extensions of natural language predicates, whereas common sense and linguistic practice scream that there are no cut-offs.

At this point, however, we could also go further in the generalization, and this is where I get my concerns. Why should the sorites consist in a series of questions in natural language? Questions are stimuli of a peculiar sort, prompting specific reactions in the form of answers. But we can imagine a sorites in which the stimuli come in a different format. Consider for instance the following variant of the scenario described by Priest. We have got a series of visual stimuli $s_0 \ldots s_n$, each in the form of a digitalized cartoon drawing. The first stimulus, s_0, is a picture of Snow White; the last stimulus, s_n, is a digital picture of a monster; and each intermediate stimulus is obtained from its predecessor by changing just one pixel. So the whole series is a short movie where Snow White is gradually warped into a monster. You are a little kid and you are sitting in front of the video screen. At the beginning you react with joy. At the end you are scared. So your initial response, r_0, is different from your terminal response, r_n. So there we are: there must be a first $k > 0$ such that $r_0 \neq r_k$. Sooner or later there must be a change—a cut-off—in the way you respond to your visual stimuli. And yet the process is perfectly gradual; the difference between any two successive frames falls below anything that is cognitively accessible to you. We are forced to recognize the existence of a cut-off point where both common sense and psychological intuition scream that there is none.

I doubt that we can say language plays any role in a scenario of *this* sort. Maybe concepts do—I'm not sure. But there is no reason to suppose that the relevant concepts are the intensions of linguistic predicates. And there are of course many such scenarios. For another example, consider a game-theoretic setup in which the series is construed as a centipede, i.e., as a game of take-it or leave-it. Money on the table accrues gradually. At the beginning of the game it is rational for you to leave the money on the table, for your opponent will do the same. At the end it is rational for you to take the money, for otherwise your opponent would take it instead. So there must be a cut-off point somewhere, a point during the game where the rational thing for you to do changes from the first type of response (leave the money) to the second (take the money). But of course the game can be construed so that the difference between any two successive stages is completely negligible for both players. For instance, we may suppose that the amount on the table accrues at the rate of just one penny each time. So there we are again. We are stuck with a cut-off that we cannot accept. And forget the fact that we can give a linguistic description of what goes on in the game. That would certainly allow us to blame it on the predicates that we use, for example on the predicate 'amount of money that a rational player (i.e., a player following the standards of ideal rationality) ought to leave on the table'. That such a predicate is ultimately vague would be an interesting result by itself.[3] But of course this is not the main story. At bottom, the problem is not one of language. It is one of rationality *tout court*.

Besides, it is not difficult to imagine scenarios in which the soritical victims are non-linguistic creatures.[4] A pigeon is trained via a regime of reward and punishment to peck to the right on the presentation of red stimuli, and to the left for other (clearly distinct) colors. Presented with a soritical series of stimuli of color patches, running from red to yellow with each patch pigeon-visually indistinguishable from its neighbor, the pigeon is bound to find itself stuck in the business of switching from right-pecking to left-pecking. Surely this has nothing to do with the vagueness of 'red', even if we could blame the trainer for scarcity of instructions.

So I think Priest is right in pointing his finger on the cut-offs. That's where we get into troubles—troubles that cannot be solved by playing around with our logical paraphernalia. But the cut-offs as such don't seem to have much to do with language, either. The troubles they originate are deeper and more general and I doubt they can be handled any better by exercising our *semantic* paraphernalia. The troubles arise at the level of cognition broadly understood. And language is but one significant part of but one sort of cognitive system.

This is not to deny that a linguistic account can be *useful*, of course. In some cases a semantic theory can be worked out that helps us understand why we are stuck with unexpected cut-off points in the presence of vague predicates, just as a rational decision theoretic account (for example) can help explain why we are stuck with cut-offs in the centipede. Of the semantic theories presently on the market Priest only mentions epistemicism and contextualism in this connection; I think supervaluationism is also in the right ball-park—whether or not it stands up well to "the cold light of inspection". I'll get back to that shortly. What I don't think is that any such theory is in the right ball-park for a general explanation of the phenomenon. In fact I am not even sure that philosophy itself is in the right ball-park for a general explanation. At bottom it may well be a phenomenon that calls for a psychological account. It may well be that cognitive

[3] See Collins and Varzi (2000).

[4] Thanks to the OUP reader for suggesting the following example.

science and the behavioral sciences at large are the only good candidates for a general explanation of the phenomenon. And it may well be that this sort of explanation is part and parcel of a general explanation of why it is that we have such a hard time dealing with the changes in our environment, when the changes are gradual. Philosophy has little to say about this—let alone semantics. (This applies to other features of the sorites paradox, too, such as the variability with which different subjects react to the same stimuli. In the Snow White scenario, for example, some children would start to get terrified as soon as Snow White's nose starts turning red while others would not get scared until the process is almost over, but it is unreasonable to suppose that such variability reflect differences in their vague idiolects or conceptual apparata. This is something that we might only be able to explain by looking into the children's individual psychophysiology. Ditto for the other cases.)

2. Semantic Explanations

Having said this, let us put the business of a general explanation aside. Let us try and focus on the limited task of assessing the explanatory force of a semantic theory—a theory designed to explain the sorites phenomenon at least inasmuch as it arises in the presence of vague predicates. Ideally, this would have to be a theory that follows from the big theory. But since the big theory is still too far away, we might as well try to work out our views bottom-up.

What options are there? As I mentioned, Priest considers two options —epistemicism and contextualism—and he has misgivings about both. On this I am inclined to agree. Concerning the first option, Sorensen and Williamson have given us a sophisticated explanation of why it is impossible to identify the semantic cut-offs of vague predicates.[5] But such an explanation does not amount to an explanation of why we are deeply disturbed by the thought that such cut-offs must exist. As Priest puts it, to explain why we cannot know the existence of something is not to explain why we find its existence counter-intuitive—and I agree with that. I also agree that the second option—contextualism—leaves the issue unresolved: the thought that there is a cut-off point in a soritical series is puzzling even if we convince ourselves that it must lie outside what Graff calls the area of "salience" set by the context[6]. I do not, however, agree that these are the only two accounts on the market to be considered in this connection. On my reckoning, supervaluationism provides a valuable alternative—in fact a better

[5] See Sorensen (1988: 189ff) and Williamson (1994), chs. 7–8.
[6] Graff (2000).

alternative. For supervaluationism does offer an explanation. Supervaluationally, the reason why we are disturbed by the existence of the cut-offs is that we cannot pin them down in any way. Ordinarily, when we recognize the truth of an existential statement we can also recognize the truth of one of its instances. We recognize the truth of a statement such as

(3) There is a number $k > 0$ such that k is greater than 3^2 but less than 4^2

because we can specify a value $k > 0$ that satisfies the condition

(4) k is greater than 3^2 but less than 4^2.

This natural impulse to demand an instance whenever an existential statement is asserted (or a counterinstance whenever a universal statement is denied) is perfectly justified in the ordinary semantics for precise languages. It is not, however, a defining feature of the meaning of the existential quantifier. At least, it is a controversial issue whether it is a defining feature; it certainly isn't in classical logic, for in classical logic existential statements are not constructive. In any event, there is no guarantee that the impulse is justified when it comes to vague languages. And this is precisely what the supervaluational account tells us. When it comes to statements asserting the existence of a cut-off point in the extension of a vague predicate, there is a tension between our natural impulse to ask for an instance, on the one hand, and the impossibility to deliver an instance, on the other.[7] We recognize the truth of a statement of the form

(1) There is a number $k > 0$ such that β_{k-1} and not β_k,

where the β_i's involve vague predicates, because we recognize that this statement is true for every precisification of those predicates. If it is true for every precisification of those predicates, then we can't go wrong if we say that it is true *simpliciter*—or so goes the story. But we cannot recognize the truth of any instance of the form

(6) β_{k-1} and not β_k

because no such instance is true for every precisification. It is true that there exists a k that marks the cut-off, but there is no k such that it is true of *it* that it marks the cut-off. This is disturbing, no question about that. Indeed it is because of this disturbing feature that many philosophers dislike the supervaluational account. ("A true existential statement must have true instances!", stamp the

[7] The tension is diagnosed in McGee and McLaughlin (1995: 207) and Keefe (2000: 185).

foot, bang the table.[8]) But I would rather say that the account deserves consideration precisely *because* of its ability to dispel this disturbing feature. Supervaluationism is in the right ball-park for an explanation of the sorites phenomenon precisely because it tells us why we are disturbed by the existence of cut-off points. We are disturbed because we can recognize the truth of an existential without being able to recognize the truth of any instance—and this goes against our natural inclination to always demand a true instance.

3. Vague Identity and Fuzzy Truth-Values

It is not my intention to insist further on the merits of supervaluationism. There may be other reasons why one could be dissatisfied with that theory—for instance, reasons having to do with the intuition that the vagueness of a predicate is part of its sense, which would make the notion of a precisification incoherent[9]—but this is not the place to take up such issues. I just wanted to illustrate my reasons for thinking that supervaluationism is in the right ball-park for a semantic explanation of the sorites. Let me now turn to Priest's own account, as presented in the second section of his paper. As I said, I believe there is a problem with this account, even if we confine ourselves to the modest task of providing an explanation of linguistic vagueness.

Let me first make a methodological remark. A great deal of Priest's account focuses on the intuition that the identity predicate is not globally transitive, and that is what I want to focus on. But Priest's working example is misleading in this regard. We are supposed to consider a sequence $c_0 \dots c_n$ of color patches such that each is phenomenologically indistinguishable from its immediate neighbors, and we are asked to agree with the soritical premise

(7) For every $k > 0$, $c_{k-1} = c_k$.

This cannot be right, for surely phenomenological indistinguishibility falls short of identity. What we want to say is that if we take the identity predicate to express phenomenological indistinguishibility, then (7) holds. Since this is a can of worms, however, I think it is better to bypass the issue altogether and work with an example that does not involve phenomenological issues, at least not explicitly. So I am going to stick to the initial example—the forced march sorites determined by the series $b_0 \dots b_n$ that begins with you, b_0, and ends in scrambled egg, b_n. Here we are stuck with a paradox insofar as the statement

[8] See e.g. Tappenden (1993: 564). But the objection is a popular one.
[9] See e.g. Fodor and Lepore (1996).

(8) For every $k > 0$, $b_{k-1} = b_k$

is *prima facie* true. And unless you have reasons to accept ontological vagueness, this means that the identity predicate is vague.

Now, Priest's account consists of three basic claims. The first, which for the sake of the argument I shall embrace, is that the vagueness of the identity predicate must be cashed out in terms of a fuzzy semantics—that is, identity statements have degrees of truth. The second claim is that the appropriate fuzzy semantics for the identity predicate is defined by the following truth conditions,

(9) $\tau\,(a = b) = 1 - d(\mathbf{a}, \mathbf{b})$,

where τ is the (continuum-valued) valuation function, d is a normalized metric, and **a**, **b** are the denotation of a, b. I am going to accept this claim, too, for (9) seems to me to be the only reasonable option for a fuzzy semantics. The third claim is that the semantic values should themselves be treated as fuzzy objects, that is, objects about which we can make fuzzy identity statements. This is actually the main idea, and it is here that we find a sketch of how a fuzzy-theoretic apparatus can provide the required explanation. It is with this sketch that I have concerns.

Take our soritical series, suppose for simplicity that $n = 9$, and consider the semantic values in the following table (which correspond to Priest's):

(*)

β_i	β_0	β_1	β_2	β_3	β_4	β_5	β_6	β_7	β_8	β_9
$\tau(\beta_i)$	1	1	1	0.8	0.6	0.4	0.2	0	0	0

In a degree-theoretic semantics we can define, besides truth, 1, and falsity, 0, a third semantic condition with respect to which a statement can be classified. We can say that a statement is *acceptable* if its truth-value does not fall below a certain threshold, ε, which in practice is determined by the context. Let us follow Priest in supposing that in the present case $\varepsilon = 0.7$. And let us follow Priest in taking the notion of a cut-off point to be defined with respect to acceptability rather than perfect truth. This amounts to saying that our responses to the relevant series of questions change in the relevant sense only when their numerical values fall below the threshold of acceptability. An acceptable identity statement is, after all, contextually acceptable even if its truth-value is not the unit. So a cut-off point is not to be thought of as the first $k > 0$ such that

(1) $\tau(\beta_k) \neq \tau(\beta_0)$

but, rather, as the first $k > 0$ such that

(11) $\tau(\beta_k) < \varepsilon$.

Taking the metric to be the standard distance $|x - y|$, it is now easy to see how the proposed explanation works. With regard to the series consisting of the fuzzy answers to our soritical questions, i.e., equivalently, the series consisting of the identity statements

(δ_i) $\tau(\beta_0) = \tau(\beta_i)$, $(0 \leq i \leq n)$

the cut-off is given by $i = 4$. But the existence of such a cut-off point is not backed up by a corresponding cut-off in the series consisting of the following identity statements,

(σ_i) $\tau(\beta_{i-1}) = \tau(\beta_i)$, $(0 < i \leq n)$

i.e., effectively, the identity statements where each answer (here: the fuzzy truth-value of β_i) is compared to its neighbors. In fact, every identity statement in this second series gets a truth-value ≥ 0.8 and is therefore acceptable, including the special case $\tau(\beta_3) = \tau(\beta_4)$. So the explanation is this: Our answers will eventually hit the cut-off; and yet, given any two neighboring answers, it is acceptable to say that they have the same truth-values. In other words, as Priest puts it, the acceptability of the β_i's must drop eventually as we go down the series, despite the fact that it never drops as we go from each β_i to its successor.

One might be tempted to object that this answer depends too heavily on the availability of a fixed value for the threshold of acceptability, ε. In order for the explanation to be satisfactory, we need an account of how the threshold value for ε is determined, and all we are told is that this value is determined by the context. But if the context determines the threshold of acceptability—the objection goes—then it also determines the cut-off between truth and falsity in a classical, bivalent setting. After all, once we settle on a value for ε we have a straightforward mapping of our fuzzy semantics into a bivalent semantics by matching acceptability with full truth:

(12) $f(\beta_i) = 1$ iff $\tau(\beta_i) \geq \varepsilon$.

So if the context does the job with respect to ε, then it should also do the job with respect to classical logic.

This objection misfires, though. For Priest's point is that the explanation does not come from the choice of ε. It comes from the fact that given any reasonable choice for ε, the truth-values of the σ_i's are all higher than ε and so these equations are all acceptable. But, of course, this property is not preserved by the map to classical bivalent semantics. With reference to our example, the equation

(13) $f(\beta_2) = f(\beta_3)$

is true (hence classically acceptable) but the next equation,

(14) $f(\beta_3) = f(\beta_4)$,

is false (and classically unacceptable).

There is another potential objection to Priest's account that misfires. As Priest observes in a footnote (n. 10), the truth-values of his fuzzy semantics should be thought of, not as precise real numbers, but as fuzzy reals—otherwise we run into a familiar objection to fuzzy logic. (Assigning a precise truth-value between 0 and 1 is tantamount to ascribing infinite precision to vague statements.) If so, then one might be tempted to protest that whether a given statement falls below the acceptability threshold may itself be indeterminate in some cases. This is not the case in our example, but that is just because the example is over-simplified. In general the (fuzzy) distance between the truth-values of the soritical statements is much less than 0.2 and therefore there may be genuine indeterminacy with regard to their acceptability. This objection is also inappropriate, though. It misfires because the relevant indeterminacy is immaterial. The counter-intuitiveness of sorites phenomena lies in the fact that there must be a cut-off, *regardless* of where exactly it is located in the soritical sequence. And Priest's explanation is that regardless of which elements of the sequence fall below the acceptability threshold, ε, any two neighboring statements in the sequence turn out to be acceptable.

So much for bad objections. But here are two worries that I believe deserves consideration. The first concerns the fact that the proposed explanation depends heavily, if not on the precise choice of ε, on the standards of precision that determine this choice. In the example under discussion, ε is set at 0.7 and the explanation goes through because each σ_i gets a value ≥ 0.8. But suppose the context was such as to determine a higher threshold, say $\varepsilon = 0.9$. In that case, only the first two and the last two σ_i's would turn out to be acceptable and the explanation would break down. In fact, we would immediately get a cut-off at $i = 3$. This would not be the same cut-off that we hit as we move along the series of the δ_i's, but that is hardly a difference that can explain why we find the latter counterintuitive. So what are we to say in cases like this?

I suppose Priest would simply say that we are dealing with an implausible scenario: the threshold is just too high. After all, in the limit case where acceptability coincides with perfect truth (i.e., when $\varepsilon = 1$) we are back to a classical bivalent scenario, where no explanation is available at all. So, more generally, the explanation may fail when acceptability is *dangerously close* to perfect truth. This may be right. But the question remains of why 0.9 (or any other value greater than 0.8, for that matter) should be regarded as a dangerous threshold in this respect. More generally, inspection shows that the proposed explanation only works when the value of ε is lower than the least value attached to the σ_i's. But I don't see how this constraint can be implemented without rendering the explanation itself dangerously circular.

The second worry is more general. As it stands, Priest's account relies on his third claim, to the effect that the basic fuzzy semantic framework must be supplemented by an account whereby the semantic values themselves are fuzzy objects. The worry is: What exactly is the cash value of this supplementary move? Suppose we don't do that. Then we could mimic Priest's explanation of what goes on as follows. Relative to the soritical series consisting of the identity statements

(β_i) $\quad b_0 = b_i \qquad\qquad (0 \leq i \leq n)$

we are forced to accept a cut-off point (in the given example: the cut-off given by $i = 4$). But relative to the series consisting of the following identities

(γ_i) $\quad b_{i-1} = b_i \qquad\qquad (0 < i \leq n)$

there is *no* cut-off point. For every such statement is assumed to be acceptable, if not plainly true. Therefore, the acceptability of the β_i's must drop eventually as we go down the series, despite the fact that the acceptability of the γ_i's never drops. End of the story.

Would *this* be a good explanation of why we find the existence of the cut-off counter-intuitive? If it were, then the detour through the meta-identity statements proposed by Priest would be unnecessary. And if it were not a good explanation then the question is what makes Priest's explanation any better.

Here is why I don't think this explanation would be a good one. It would not be a good explanation because it would simply amount to a redescription of the puzzle. After all, the sorites phenomenon is puzzling precisely because the difference between any two neighbors of the given sequence is negligible whereas the difference between the first member and some later, distant member is not negligible. Small differences accrue—that's exactly the problem. Of course a fuzzy logician can infer from this that the identity predicate is not globally transitive, but that doesn't add much. Unless we tell a story that explains why this outcome is at odds with the intuition that identity is transitive, the claim that identity is not globally transitive simply reports the facts. I am sure Priest agrees with this, for otherwise he would have thought that his earlier attacks on the problem already had the explanatory force that he is after. (On such earlier accounts, the explanation traded on the difference between local and global validity. We mistakenly think that inferences that can be used over short distances—such as inferences exploiting the transitivity of identity—are reliable over long distances as well, which is not the case: in particular, identity is transitive locally but not globally.[10])

So if this simplified explanation is not good, what exactly do we gain by making the extra step of treating the semantic values themselves as fuzzy objects? As far as I can see we don't gain much. It is true that the γ_i's and the σ_i's are distinct: in one case we have a series of identities concerning pairs of neighboring b_i's (the objects obtained by replacing i molecules of your body with i molecules of scrambled egg); in the second case, corresponding to Priest's strategy, the series of identities concern pairs of neighboring β_i's (which are statements to the effect that you are identical with the object obtained by replacing i molecules of your body with i molecules of scrambled egg). And whereas we may suppose that the identities of the first series are all perfectly true, the identities of the second series need not be true: what matters is that they are acceptable. However, such differences are irrelevant when it comes to the requested explanation. We are going from one soritical series to another, and what we are told about the second case is exactly what the simplified account tells us about the first: in both circumstances we are dealing with a series of neighbor-identities (the γ_i's and the σ_i's, respectively) all of which are acceptable despite the fact that the corresponding distance-identities (the β_i's and the δ_i's) eventually drop below the acceptability threshold. To put it differently, Priest's semantics satisfies the following biconditionals:

(15) β_i is acceptable iff δ_i is acceptable $(0 \leq i \leq n)$
(16) γ_i is acceptable iff σ_i is acceptable $(0 < i \leq n)$

i.e., effectively:

(15') $b_0 = b_i$ is acceptable iff $\tau(\beta_0) = \tau(\beta_i)$ is acceptable $(0 \leq i \leq n)$
(16') $b_{i-1} = b_i$ is acceptable iff $\tau(\beta_{i-1}) = \tau(\beta_i)$ is acceptable $(0 < i \leq n)$

And since acceptability is what matters, it seems to me that this makes the detour proposed by Priest redundant. If it is redundant, it adds no extra value. If it adds no extra value, then the explanation that we get by taking the detour cannot be any better than the explanation that we get without taking the detour.

[10] See e.g. Priest (1998).

And as we have seen, the explanation that we get without taking the detour is not good enough.

4. Conclusion

My conclusion, therefore, is not too optimistic. I don't think the novel account offered by Priest will take us closer to the heart of the problem. On my reckoning, the supervaluational account would still fare better, but never mind that. At this point I would rather be inclined to end on a note of methodology. We have got many semantic accounts. None of them seems capable of delivering the whole story, but many succeed in diagnosing at least some of the peculiarities of the sorites phenomenon. So perhaps a good thing to do at this point would be to try and put the pieces together. On the other hand, because of my earlier remarks about the non-linguistic nature of the phenomenon in its most general form, I am not too optimistic that this strategy will take us very far, either. The sorites is a deep and bewildering puzzle precisely because it arises at a deep and fundamental level, one that appears to be prior to the engagement of any logical *and* semantic paraphernalia. Unless we take that into serious account, I doubt we can achieve the sort of explanation Priest is after, which is why I think we are still stuck with the puzzle. In *Herzog*, Saul Bellow tells a story about a club in New York that most of us should find familiar.[11] It's that club where people are the most of every type. There is the hairiest bald man and the baldest hairy man; the shortest giant and the tallest dwarf; the smartest idiot and the stupidest wise man. And what do they do? On Saturday night they have a party. Then they have a contest. And if you can tell the hairiest bald man from the baldest hairy man—says Herzog—you get a prize. My impression is that if we entered the contest we might perform better than our fellow non-philosophers. In the end of the day, however, I suspect we would still be puzzled. We would still find ourselves looking at the prize with an overwhelming sense of incredulity—that sense of incredulity that we experience whenever we hit a cut-off.[12]

References

Bellow, S. (1964), *Herzog*, New York: Viking Press.
Collins, J. D., and Varzi, A. C. (2000) 'Unsharpenable Vagueness', *Philosophical Topics* 28: 1–10.

[11] See Bellow (1964: 295–296).
[12] Thanks to Matthew Slater for helpful exchanges on the topic of this paper.

Fodor, J. A., and LePore, E. (1996), ‘What Cannot Be Evaluated Cannot Be Evaluated, and It Cannot Be Supervalued Either’, *Journal of Philosophy* 93, 516–535.
Graff, D. (2000), ‘Shifting Sands: An Interest-Relative Theory of Vagueness’, *Philosophical Topics* 28, 45–81.
Keefe, R. (2000), *Theories of Vagueness*, Cambridge: Cambridge University Press.
McGee, V., and McLaughlin, B., 1995, ‘Distinctions Without a Difference’, *Southern Journal of Philosophy* 33 (Suppl.): 203–252.
Priest, G. (1998), ‘Fuzzy Identity and Local Validity’, *The Monist* 81: 331–342.
Priest, G. (2003), ‘A Site for Sorites’, this volume.
Sorensen, R. A. (1988), *Blindspots*, Oxford: Clarendon Press.
Tappenden, J. (1993), ‘The Liar and Sorites Paradoxes: Toward a Unified Treatment’, *Journal of Philosophy* 90: 551–577.
Williamson, T. (1994), *Vagueness*, London: Routledge.