Draft of paper for Dummett LLP volume
NOT FOR OTHER CIRCULATION

CRISPIN WRIGHT

"Wang's Paradox"

I Introduction

There is now a widespread accord among philosophers that the vagueness of natural language gives rise to some particularly deep and perplexing problems and paradoxes. It was not always so. For most of the first century of analytical philosophy, vagueness was generally regarded as a marginal, slightly irritating phenomenon, —receiving some attention, to be sure, in parts of the *Philosophical Investigations* and in the amateur linguistics enjoyed by philosophers in Oxford in the 1950s, but best idealised away in any serious theoretical treatment of meaning, understanding and valid inference. Frege, as is well known, had come to be thoroughly mistrustful of vagueness, supposing that a language fit for the purpose of articulating scientific and mathematical knowledge would have to be purified of it. Later trends in philosophical logic and semantics followed his lead, not indeed in setting about the (futile) task of expurgating vagueness from natural language but by largely restricting theoretical attention to artificial languages in whose workings vagueness was assigned no role.

During the 1970s this broadly Fregean disdain for vagueness was completely turned about. The thirty years since have seen a huge upsurge in interest in the topic and publications about it, most of them by philosophers with not one iota of sympathy with the approach of "Ordinary Language Philosophy" or the seemingly haphazard and anti-theoretical remarks of Wittgenstein. The reasons for the sea-change are no doubt complex but my own belief is that a crucial impetus was provided by a single publication: the 1975 special number of *Synthese*[1] in which a number of subsequently influential papers were published for the first time but in which the single most important contribution — the paper one would recommend to a philosopher who was only ever going to read one essay on the topic— was Michael Dummett's "Wang's Paradox".

Dummett's paper was actually written some five years earlier and had already had a significant degree of circulation.[2] It focuses only partially on vagueness, giving as much or more attention to strict finitism as a rival to the intuitionist philosophy of mathematics — arguably, indeed, as the proper logical conclusion of the intuitionistic tendency — and to the nature of observational language. The impact of the paper was in no small measure due to

---

[1] *On the Logic and Semantics of Vagueness, Synthese* 30, Nos. 3/4.

[2] It was the single most important influence on my own first study of the topic, "On the Coherence of Vague Predicates", and also I believe on Kit Fine's prototypical supervaluational treatment of the semantics of vagueness, "Vagueness, Truth and Logic", both of which were first published in the same number of *Synthese*.

these simultaneous concerns: to the connections it made between topics — vagueness, observationality, and finitism in the philosophy of mathematics — which had received no systematic concerted discussion before. But it also contained a strikingly clear display, unmatched in any previous discussion of which I am aware, of the simple essential architecture of one form of the Sorites paradox. The form of the paradox in question involves an observational predicate, F,[3] and a finite series of items connecting a first F-element with a final non-F element, but with each element relevantly indistinguishable (at least by unaided observation) from those immediately adjacent to it. Such a series is possible, of course, because and only because observational indistinguishability is not a transitive relation. Since the thought is, at first blush, utterly compelling that an observational predicate cannot discriminate between observationally indistinguishable items, it seems we have to accept that anything in the series that is F is adjacent only to things which are also F (call this principle the Major Premise.) And that, on the stated assumptions, is enough for paradox. I think some foggy notion had prevailed earlier that the antinomy was somehow due to applying to vague expressions principles of reasoning appropriate only to precise ones. Dummett seems to have been the first explicitly to register simultaneously both the *utterly plausible* character of the major premises in a very wide range of examples and the *utterly basic* character of the logic required: if vague expressions are not fit for reasoning involving merely iterated applications of *modus ponens* and universal instantiation, it is hard to see how they can be fit for reasoning at all. I do not think that the depth of the crisis for common sense which this paradox involves had really been properly appreciated before Dummett's discussion.

Dummett's own reaction was dramatic and is worth quoting in some detail. What the paradox should teach us, he writes, is that:

> ..... the use of vague predicates — at least when the source of the vagueness is the non-transitivity of a relation of non-discriminable difference — is *intrinsically incoherent* [my italics][4]

He then elaborates this conclusion in four subsidiary claims:

> (1) Where non-discriminable difference is non-transitive, observational predicates are necessarily vague
> (2) Moreover, in this case, the use of such predicates is intrinsically inconsistent
> (3) Wang's paradox merely reflects this inconsistency. What is in error is not the principles of reasoning involved, nor, as on our earlier diagnosis, the [major premise]. The [major premise] is correct, according to the rules of use governing vague predicates such as 'small'; but these rules are themselves inconsistent, and hence the paradox. Our earlier

---

[3] Dummett characterises an observational predicate as one "whose application can be decided merely by the employment of our sense organs", "Wang's Paradox" p. 320.

[4] "Wang's Paradox" p. 319

— [proto-supervaluationist] —

> model for the logic of vague expressions thus becomes useless: there can be no coherent such logic.[5]

The fourth subsidiary conclusion, in keeping with (2), then dismisses as correspondingly incoherent the conception of mathematical totalities as the extensions of vague predicates advanced by strict finitism.

On first encounter, Dummett's principal conclusion, that vagueness infects natural language with inconsistency, seems desperate. And there would be, to be sure, more than a suspicion of *non-sequitur* in the transition from an argument that observational expressions have to be both vague and governed by inconsistent rules to the conclusion that vague expressions *per se* are governed by inconsistent rules. Observationality, understood in a way that suffices for the major premise, entails both vagueness — an observational predicate will fail to draw a line between indistinguishables — and inconsistency, at least whenever Fs can be ancestrally linked with non-Fs via a chain of indistinguishable pairs. That suffices for the conclusion that vague predicates are one and all governed by inconsistent rules only if all vague predicates are observational. And that is not true — recall the strict finitist's predicates of practical intellectual possibility — "intelligible" as applied to numerals, "surveyable" as applied to proofs, and so on — and indeed "small" as applied to numbers. Moreover even in observational cases, it is apparently the observationality itself that directly generates the problem, rather than the lack of sharp boundaries that it enjoins. Dummett, that is to say, has described a paradox of *observationality* and — it is tempting to say — its solution must consist in an improved understanding of what an expression's possessing observational content involves. But he has made no case for saying that *vagueness* is intrinsically paradoxical — which is just as well when one considers that, whether or not there are any purely observational expressions in the sense the paradox exploits, vagueness is the norm for expressions of natural language.

However a little reflection shows that the concern does indeed ramify across vague expressions as a class. It is constitutive of an expression's being vague, surely — or so one might think — that it should fail to draw a sharp boundary in a suitable Sorites series. Consider a column of soldiers marching past their commanding officer, ranging from five feet six to six feet six inches in height and so lined-up that each man is marginally shorter than the man who immediately succeeds him. The *precision* of the predicate, "is more than six feet tall", consists in the fact that, no matter how small the differences in height between one man and the next, there is certain to be a first man to which it applies. Correspondingly, since vagueness is the complement of precision, the vagueness of "short" should consist in

---

[5] "Wang's Paradox" p. 319-20

the fact that, on the contrary, if the differences in height are sufficiently marginal, there need be no sharp bound on the short men in the march-past: no last man who is short followed by a first man who is not. But if there is no such boundary, then whenever a man is not short, he cannot be immediately preceded by a short man. Since all the men who are not short come relatively late in the march-past and hence do have predecessors, the latter likewise cannot be short. Lack of sharp boundaries *as such* thus does seem to imply paradox. To say that F lacks sharp boundaries in a series of the germane kind is to say, it seems, that there is no element, x, which is F but whose immediate successor, x', is not. That is a claim of the form,

$\sim(\exists x)(Fx \ \& \sim Fx')$

and is accordingly classically equivalent to the major premise

$(\forall x)(Fx \rightarrow Fx')$

for exactly the kind of Sorites that Dummett focused on. In brief, in the presence of classical logic, vagueness apparently *consists in* the holding of the major premise for the Dummettian Sorites. For a predicate to lack sharp boundaries does indeed imply, in that setting, that it is subject to inconsistent rules of application.

Worse, this form of the Sorites paradox — the *No Sharp Boundaries paradox* — survives even if we drop the classical logical setting. Where k ancestrally succeeds 0, the logic required for eliciting paradox from the trio, $\{F0, \sim Fk, \sim(\exists x)(Fx \ \& \sim Fx')\}$ needs nothing to allow conversion between the quantifiers but consists merely in the standard introduction rules for conjunction, for negation (i.e., *Reductio ad Absurdum*), and for '$\exists$'. These principles are intuitively no less constitutive of the content of the constants they govern than are universal instantiation and *modus ponens*. So there is a powerful-seeming line of argument that appears to drive us towards Dummett's — as it appeared initially, over-stated — conclusion. Vagueness *per se* does appear to infect natural language with inconsistency.

David Pears once remarked that the characteristic effect of a Dummett intervention in a philosophical conversation was as if to turn on a light which the others had overlooked in a gloomy room. Sometimes, of course, what better lighting shows up is not the solution to a problem but its real contours. That was the kind of illumination shed by "Wang's Paradox". My own thinking about vagueness and finitism and the associated cluster of issues whose connections Michael's paper displayed had benefited from discussions with him going back almost a decade before its publication to when I was a PhD student. My hope is that, in this distinguished volume to debate and celebrate his wonderful contributions to modern philosophy, he will enjoy the spectacle of me still wrestling, almost forty years on, with the same conundrum.

II Vagueness as Semantic Incompleteness

Frege's disparaging view of vagueness is seldom explicitly argued for in his writings but the

little he says does indeed suggest he thought that the phenomenon threatens the stability of basic logic. However he does not cite the Sorites paradox in support of this complaint. Frege takes it that if a term is vague, that will be tantamount to its being only partially defined — so that it will only be of things of a certain kind that it will make sense to say that it either applies to them or that it does not. But that will give rise to failures not just of excluded middle, but of other laws too— contraposition, for example. Everything to which the predicate applies will be a thing of the presupposed kind; but it will not be correct to affirm, conversely, that everything not of that kind is something to which the predicate doesn't apply — since the range of cases of which the predicate may significantly be denied is likewise restricted to the kind of thing in question.[6]

The most interesting thing about this line of thought is not its conclusion, but its premise: the equation of vagueness with partial definition. To fix ideas, consider an artificial example of Tim Williamson's: the predicate, 'dommal', whose satisfaction conditions are stipulated as met by a creature which is a dog, and failed by a creature which is not a mammal.[7] Nothing else is said by way of determination of its meaning, so the effect is that it is undefined whether 'dommal' applies to mammals other than canines. The conception of vagueness in ordinary language to which Frege gives one of the earliest expressions is in effect exactly that it is in general the naturally occurring counterpart of the artificially generated indeterminacy of 'dommal'. Our training gives us rules for the application of 'bald'— roughly, looking sufficiently like certain paradigms — rules for applying 'not bald' — again, looking sufficiently like certain paradigms — and these rules once again fail between them to cater for all cases. The exceptions — the borderline cases — are those for which we lack any sufficient instruction. They stand to 'bald' essentially as non-canine mammals stand to 'dommal'.

This idea conditioned virtually all mainstream work on vagueness until quite recently. It has been so widely accepted as to be either unnoticed or received as a datum of the problem. So conceived, vagueness is a matter of *semantic incompleteness*. A vague expression is one for which we have mastered rules for assenting to its application and rules for denying it which between them leave space for a *gap* — a range of cases where the rules simply don't give us any instruction what to do.

I do not know whether this way of thinking about vagueness implicitly originated

---

[6] I take thisto be the argument of *Grundgesetze*, vol. I, §65. Frege does mention the Heap itself at *Begriffschrift* _27, in the context of his definition of what it is for a property to be hereditary in a series, but remarks on no threat to logic, suggesting merely that, since 'heap' is not everywhere sharply defined, we may regard the major premise as indeterminate. (That is no option of course if we take it that the minor premise, that 0 grains of sand cannot constitute a heap, and the conclusion, that—say—200,000 grains of sand cannot constitute a heap, are, respectively, true and false. An indeterminate statement cannot be inconsistent with the facts! Even Frege, it seems, underestimated the Sorites.)

[7] Timothy Williamson *Identity and Discrimination*, (Oxford: Blackwell 1990) at p. 107.

with Frege, or whether it is much older. In any case it is, undeniably, extremely natural.[8] To be sure, not all writers about vagueness nowadays think of it as a semantic phenomenon at all, but among the majority who still do, the Fregean conception is entrenched. It is, for example, a presupposition of the whole idea that vague expressions allow of a variety of alternative but admissible sharpenings, whose effect will be to include or exclude certain items from their range of application whose status was previously indeterminate. This idea in turn, of course, is presupposed by the widely accepted supervaluational approaches to the semantics and logic of vague discourse.

Natural as it may be, however, the Fregean conception now seems to me to be almost certainly wrong. At the least, it is open to serious objection on two major counts. First, it is a very poor predictor of our actual linguistic practices. It gives the wrong prediction about our responses to — and responses to responses to — borderline cases of standard Sorites-prone predicates. Someone who has mastered 'dommal' will know better than to apply it, or its contrary, to non-canine mammals. Asked if a cat is a 'dommal', he will, or should, say that he is not empowered to judge — for all there is to go on is a sufficient condition for being a 'dommal' and a necessary condition for being a 'dommal', and cats neither pass the first nor fail the second. By contrast, what we find in borderline cases of the distinction between, say, men who are bald and men who aren't it is exactly *not* a general recognition that there is no competent verdict to return but rather a phenomenon — spreading both among the opinions of normally competent judges and, across time, among the opinions of single competent judge — of weak but conflicting opinions, unstable opinions and — between different judges — agreement to differ. It is true that sometimes a competent judge may simply be unable to come to a view but it is not a necessary characteristic of the borderline case region that it comprises just — or even any — cases where competent judges *agree* in failing to come to a view; and any case about which a competent judge fails to come to a view may, without compromising his competence, provoke a (weak) positive or negative response from him on another occasion. Moreover, failure to come to a view is not the same as the judgement that there is no competent view to take — and it is the latter that is appropriately made of a cat by someone competent with the use of 'dommal'.

These considerations are radically at odds with what the Fregean conception would lead one to expect. If someone takes the view that some particular cat is a dommal, then, *ceteris paribus*, they show that they have misunderstood the explanation of the word. If someone understands the explanation properly, they won't return a verdict about a cat. In contrast, our responses to those who do return verdicts in the borderline area of 'bald' is that, provided those verdicts are suitably sensitive and qualified, it is permissible so to do. We are *liberal* about judgements in borderline cases. Our thought is not that they are cases about

---

[8] Thus Kit Fine, op. cit., p. 265: "I take [vagueness] to be a semantic notion. Very roughly, vagueness is deficiency in meaning."

which one ought to have no view but rather merely that they are cases about which it's probably pointless to try to work through differences of opinion. The psychologist testing the responses of a variety of, by normal criteria, competent subjects down a Sorites series would *expect* divergences in the borderline area — indeed, that is what the borderline area *is*: an area of expectable and admissible divergence. If the 'dommal' model were correct, the expectation would be of consensual silence.

The second major area of difficulty for the Fregean conception concerns one of the most arresting and disconcerting features associated with the kind of vagueness that interests us, that of so called *higher order* vagueness. There are various ways of eliciting the impression that the phenomenon is real and needs to be reckoned with. One line of thought is outlined by Dummett:[9] if I was to introduce a new word —say 'sparsey' — to apply to just to borderline bald men, we would find that the boundary between the bald men and the sparsey men was itself vague. Another line of thought is to observe that, in a typical Sorites series, there will be no determinate first case of which we are content to judge that doing *something other* than returning the initial positive verdict is appropriate. How is this to be explained under the aegis of Fregean conception? According to the Fregean conception, borderline cases are cases where we have not provided for a verdict — cases that we failed to cover by the relevant semantic rules; so the remaining cases are ones where we *have* so provided and a negative or positive version is appropriate. How can *this* distinction in turn be one for the drawing of which we could somehow have made insufficient provision? No doubt it is up to us what provision we've made — Williamson might for example have provided even less for 'dommal', restricting the sufficient condition to Corgis. But whatever provision is made, it shouldn't then need a *further* provision to settle which cases it does and doesn't cover respectively. Giving just the provision that he did, Williamson thereby settled that the borderline cases of 'dommal' comprise all non-canine mammals. He didn't merely settle that *some* kinds of case are to be borderline for 'dommal', leaving it open whether others are borderline or are cases where a determinate verdict is mandated. That matter was determined for him by the nature of his omission. It did not need further determination by him — indeed, it was not for him to stipulate at all — how far the omission extends.

Of course a defender of the Fregean conception has possible responses to this. It is intelligible, for instance, how an expression with a semantic architecture like that of 'dommal' might nevertheless allow vagueness on the boundary between the things which satisfied its sufficient condition and the things that neither satisfied its sufficient condition nor failed its necessary one. These — second-order — borderline cases may precisely be things for which it was indeterminate whether or not they satisfied the relevant sufficient condition. Such indeterminacy could arise because the concept giving that sufficient

---

[9] In "Wittgenstein's Philosophy of Mathematics" *Philosophical Review*LXVIII (1959), 324-48, (reprinted in M. Dummett *Truth and Other Enigmas* (London: Duckworth 1978); the cited line of thought occurs at p. 182.)

condition — *dog* in the Williamson example — might *itself* have the same kind of semantic architecture. Similar possibilities would apply to the concept — *mammal*— that gave the original necessary condition. Since there seems to be no limit to the extent to which the pattern might be iterated, it looks as though the Fregean conception might after all be able to recover some of our preconceptions about higher order vagueness, to whatever extent they run, provided the relevant concepts that would be successively invoked are all of the illustrated semantic-architectural kind.

That is a proposal. But it seems to have very little mileage in it. When we reflect on the prototypical Sorites-prone predicates — predicates like 'red', 'short', 'bald', 'heap,' and so on — the most salient feature about them is their *immediacy*: it simply isn't credible that our conception of their conditions of proper application is informed by an indefinitely extending structure of partial definitions, each one deploying novel concepts distinct from those employed in the sufficient, or necessary, conditions articulated by its predecessor.

These are serious difficulties. But higher order vagueness poses a further and I think decisive problem for the Fregean conception. Simply: it allows of no coherent description in terms of the Fregean template for what a borderline case is.[10] The borderline cases of 'dommal' are cases where there is mandate neither to apply 'dommal' nor to apply is contrary. Borderline cases of borderline cases, on this model, will thus be cases where there is neither mandate to apply a predicate or to apply its contrary, nor mandate to regard them as borderline — to regard them as cases where neither the predicate nor its contrary is mandated to apply. So they are cases where there is no mandate to apply the original predicate (otherwise they wouldn't be borderline cases at all), no mandate to apply its contrary (for the same reason), but also no mandate to characterise as cases where there is neither mandate to apply the predicate nor mandate to apply its contrary! That's absurd. Whenever— as the first two conjuncts say — a case is such that there is no mandate to apply the predicate to it and no mandate to apply its contrary, then it will be *true* — and hence mandated — to say just that; but that is exactly what the third conjunct denies.

In sum: the Fregean conception — the conception of vagueness as semantic incompleteness — is in tension with our actual sometimes positive or negative reactions to borderline cases, at odds with the liberality of our reactions to others' reactions to them, and, on two counts, can provide no room for the — at least apparent — phenomenon of higher order vagueness.

It may be rejoined in mitigation that we should by now have learned to expect that *any* possible broad conception of vagueness will have problems in accommodating one or another of our preconceptions about the matter, or in respecting certain aspects of the phenomenology of vagueness or our linguistic practice with vague expressions. That, indeed,

---

[10] See C. Wright,""Vagueness: a Fifth Column Approach" in J.C Beall and Michael Glanzberg, eds., *Liars & Heaps: New Essays on Paradox* (Oxford: Oxford University press, 2003) at p. 89.

is what makes the whole issue so hard. Even if resolute prosecution of the Fregean conception is going to require a theory which is indeed in tension with certain aspects of the apparent data, that — it may be said — should be a decisive consideration against it only if some other theory can match the available advantages of such a theory, including those made possible by the apparatus of supervaluation, whilst causing fewer, or less serious casualties. But I disagree. I think the problems I have just outlined are sufficiently serious to justify persisting in the assumption that Fregean conception, however natural, is mistaken: borderline cases of vague expressions of the type typified by the 'usual suspects'— "red", "bald", "heap", "short", and so on — simply are not to be conceived in terms of the idea of, as it were, one's semantic instructions giving out, of the rules of language failing to provide guidance. But if they are not to be so conceived, what are the alternatives?

III Vagueness as Unknown Precision

If the borderline cases of a vague expression are not to be viewed as cases for which its governing rules fail to prescribe any verdict, then there seem to be just two remaining possibilities: either the rules are inconsistent — Dummett's idea— and thus prescribe conflicting verdicts in every case, or we should swallow hard and accept that *consistent* verdicts are indeed prescribed in every case. In any Sorites series of the normal 'monotonic' kind, the latter is tantamount to an acceptance that the rules governing the affected predicate mandate a *sharp cut-off*. Bivalence is therefore assured, and with it classical logic. What is missing, though, is any account of what vagueness is or why it arises.

As we know from the work of Williamson and others,[11] this can be made to be a much more resilient proposal than at first appears. Its proponent — the *epistemicist* — is thinking of our use of vague expressions as broadly comparable to the sorts of judgements we should make about the application of precise predicates like 'is more than six feet tall' if excluded from reliance on canonical means of measurement and forced to estimate by eye. There would be cases where it would be clear that a subject was more than six feet tall, cases where it was clear that he was not, and cases where observation left one unable to form a view, or where our views were weak and unstable. This model thus predicts something very like our actual patterns of application of vague expressions in borderline cases. Presumably these patterns of reaction would be picked up by someone who was trained in the use of 'is more than six feet tall' purely by immersion in the practice of those restricted to visual estimation and to whom no communication was made of the real meaning of the predicate, in virtue of which its extension is actually sharp. We can construct a fantasy around that idea.

---

[11] The leading systematic account of the epistemicist view is Timothy Williamson's *Vagueness* (London: Routledge 1994.) See also ch. 6 of Roy Sorensen's *Blindspots* (Oxford: Oxford University Press 1988.)

Call those privy to the real meaning the *Priests* and those kept in the dark the *Acolytes*. We might imagine that, as time goes by, the Priests die out but the practice of the Acolytes survives in more or less stable form, though uninformed by any adequate conception of what determines the extension of 'morensicksfittle' as the phrase begins to be written in their dialect.

I do not know if Williamson, or any other defender of the epistemic conception, would take any comfort in this parable. Certainly, it is no part of their view that we should think of ourselves as having, over generations, lost touch with earlier fully explicit conceptions of the meanings of vague expressions. What cannot be avoided, however, is the admission that, if the epistemicist view is right, we do not as a matter of fact *have* satisfactory such conceptions and have no idea how to go about recovering them. This is something that epistemicism needs to explain, and — so far as I am aware — no explanation has ever been offered. It's true that Williamson, for one, has worked hard to explain why we cannot know where the (postulated) sharp cut-offs come in Sorites series.[12] But it is not an explanation of that that's being asked for. Estimations of the application of 'is more than six feet tall' based on unaided observation would indeed be subject to margins of error, and for my present purposes we can concede that Williamson explains why, accordingly, we cannot know by such means where, in a march-past of fifty men beginning at five foot six inches tall and increasing by small variable margins less than one quarter of an inch each time, the first man more than six feet tall is to be found. But that's not to explain why I cannot know the alleged principle or principles that stand to my use of 'bald' as the condition expressed by the predicate, 'is more than six foot tall' stands to the practice — at least according to the suggestion of the parable — of the Acolytes.

The real difficulty, though, is to make anything of the suggestion that the Acolytes' practice is indeed actually so informed. It is only because the Priests understood what it is to be more than six feet tall that they had any inkling about when it is safe to apply the predicate, or to deny its application, under circumstances of unaided observation. The practice so informed may be imitable in a relatively stable and transmissible way, but insofar as it is, there is no longer any motivation for saying that it is driven by the original precise principle. What drives it, rather, is a sense of the patterns of use — in which, for the Priests, the precise principle was once instrumental — purely as patterns *in their own right*. There is no sense in the idea of a continuing additional undertow exerted by an originating principle which nobody any longer grasps. The most fundamental difficulty with the epistemicist proposal is not merely — to put it unkindly — the element of superstition involved in the suggestion that there are indeed, in the case of all vague expressions which contribute to a successful linguistic practice, underlying principles which determine sharp extensions. The

---

[12] See especially ch. 8 of his *Vagueness*.

charge of superstition might be thought to be addressed, at least to a degree, by the reflection that what at this point may appear to be the only possible alternatives — semantic incompleteness, a la Frege, and incoherentism, a la Dummett — also appear radically unsatisfactory But the more acute problem is that, so long as the epistemicist has to concede that we have no inkling of what the relevant principles are or how they might be ascertained, it is merely bad philosophy of mind to suppose that our linguistic practice consists, in some sense, in their implementation — that it is, in any meaningful sense, *regulated* by them at all.

IV Vagueness as Incoherence

If both Incompleteness and Epistemicism are unacceptable, then rather than merely dismiss it as "radically unsatisfactory", we should reconsider Dummett's incoherentist response before going any further. The leading thought is that the rules governing a vague expression do indeed provide guidance right through Sorites series, rather as a powerful river current guides items of flotsam over the waterfall! The idea that, in our mastery of natural language, we are governed by inconsistent rules is counter-intuitive to be sure. But is there some more fundamental objection? A broadly successful practice *can* be informed by incoherent rules: think of simple division taught to children without any explicit proscription of division by zero, or the game of croquet which, I am told, was for years codified by official rules which contained an inconsistency about the permissibility of iterated roquets. Why should it not be so with linguistic competence?

Well, one crucial question is what explanation the incoherentist can offer of our characteristic patterns of response to borderline cases. Why in a Sorites series do we start out with strong positive opinions, then gradually lapse into weak, defeasible opinions, conflicting opinions, unstable opinions, and failures to come to an opinion, before finally gradually reverting to strong negative opinions? Why doesn't confusion reign throughout the range? The incoherentist may try to suggest that the deterioration of our linguistic practice from broad consensus into paralysis and conflict is a function of our realisation that it is to disaster that the rules are taking us. But that does not explain the subsequent stabilisation of confidence in negative verdicts. And it is in any case a poor explanation of the distinctive patterns of reaction in the borderline area. It's a poor explanation of those patterns for the simple reason that it doesn't need the context of a Sorites paradox to elicit the characteristic responses involved: they are elicited anyway by confrontation with borderline cases, from thinkers who have no inkling of the Sorites or even — in the case of, say, young children — any capacity to follow the reasoning of the paradox and see its point.

There is more. The suggestion that in our linguistic practice with vague expressions we follow incoherent rules —rules that do indeed actually mandate the application both of an affected predicate and of its contrary to the very same object— is powerless to explain the

basic point that, when we do confront a Sorites paradox, we have *not the slightest inclination* to weigh the two limbs of the contradiction equally. There is absolutely no inclination to regard the verdict reached by the Sorites chain as correct. It is utterly dominated by the verdict with which it conflicts, and the paradox initially impresses as the merest trick. By contrast, where we really do have conflict in the rules governing a concept — for example, the concept of *course of values* introduced by Basic Law V of Frege's *Grundgesetze* — the two components of the paradox are balanced in our esteem, and we have not the slightest sense that one is to be preferred to the other. This is a basic datum which any satisfactory account of vagueness should accommodate and explain, but which — at least on the face of it — the incoherentist view is powerless to explain.

As observed, it is not to be denied that sense can be made of an in certain respects successful activity being informed by inconsistent rules. The society that teaches rules of division which are correct except for the detail that they allow division by zero, may get away with it because it never occurs to anyone to attempt a calculation involving a ratio with zero as its denominator. Or it may be that practitioners are aware of the contradiction, but don't exploit it and manage to avoid situations in which it matters. In any such case, though, the warrant to identify contradiction in the rules is wholly dependent on the way they are officially *explicitly* codified: the character of the informed practice, considered just in its own right, furnishes — and in so far as it is stable, successful and communicable, can furnish — *no* grounds to propose a theory of it which represents its rules as inconsistent. The point is surely equally good for our pervasive and remarkably successful linguistic commerce involving vague expressions. To represent it as the product of inconsistent rules purely on the basis of a paradox which no-body actually accepts provides not merely for poor explanations of linguistic practice, in the respects just noticed, but is empirically entirely unmotivated.

V Vague Discourse as Unprincipled

Now, though, we appear to have hit a complete impasse. What are we to think of the rules which govern the use of a vague expression as instructing us to do as we advance down a Sorites series into the borderline area? It seems that there are just three possibilities: either there's no prescription in the borderline area, or the prescription remains the same as it was, or it changes abruptly — semantic incompleteness, incoherentism, and (in effect) epistemicism. Yet we have reviewed serious causes for discontent with each of these three proposals. Each, indeed, comes short in the most basic way, by failing to offer a satisfactory explanation of one or another aspect of competent speakers' linguistic practice with vague expressions. Semantic incompleteness fails to explain, for instance, our tolerance of conflicting verdicts in the borderline region; incoherentism fails to explain, among other things, the disparity in our reaction to the two components of the contradiction; and

epistemicism, in so far as it is content to appeal to underlying semantic features transcending any sense that competent speakers have of proper linguistic performance, seems to disclaim any ambition of giving an explanation of our linguistic practice with vague expressions at all.

The puzzle is intense. But we get a pointer to what I've long believed to be the direction in which to find the correct response to it if we consider the so-called 'Forced March' variation on the Sorites paradox. The Forced March involves no reasoning to a contradiction from premises, plausible or otherwise. Rather we simply take a hapless subject case-by-case down a Sorites series, ranging from things that are clearly F at one end to things that are clearly not F at the other, and demand a verdict at every point. The subject starts out, naturally, with the verdict, F. And there are just two possibilities for what happens afterwards. Either he goes on returning that verdict all the way down, or at some point he does something different — if only refusing to issue a verdict at all. If he does the former, then eventually he'll say something false, and hence will betray *incompetence*. And if he does the latter, then he will draw a distinction by his responses which (i) will have no force of precedent for verdicts in other contexts, and (ii) will correspond to no relevant distinction that he can call attention to between the last case where he gave the original verdict and the first case where he changes. So his verdicts will be *unprincipled*. Conclusion: anyone who uses a Sorites-prone expression can be forced to use it in ways that are either incompetent or unprincipled. If we add the plausible-seeming supposition that competent linguistic practice is always essentially principled — always consists in the proper observance of semantic and grammatical rules — then the conclusion, on either horn, is that the use of vague language is bound to be incompetent.

The solution must be to break the tie between 'unprincipled' and 'incompetent'. But that's to say there has to be sense in which competent classification utilising vague expressions does not consist in the implementation of the requirements of semantic rules. When the question is, what do the semantic rules subservience to which constitutes competence for a vague expression require when it comes to borderline cases, the answer we should give is not *any* of the three canvassed. The position is neither that the requirements give out, nor that they remain in force driving us on towards paradox, nor that they mandate a sharp cut-off (of some kind). Rather, it is that competence with basic vague expressions is *not a matter of subservience to the requirements of rule at all.*

This is, indeed, the conclusion to which I came in my paper on the Sorites published in the same volume as "Wang's Paradox".[13] The great difficulty now, as then, is to understand what we are committing to if we accept it —and is the more acute when one reflects that, on at least one way of understanding the issue, to claim that basic classifications effected using

---

[13] C. Wright *op. cit.* n. 2 and "Language-Mastery and the Sorites Paradox" in G. Evans and J. McDowell, eds., *Truth and Meaning*, (Oxford: Oxford University Press 1976) 223-247. (Reprinted in R. Keefe and P. Smith, eds., *Vagueness: a Reader*, Cambridge, Mass.: Bradford/MIT. 1996.)

Sorites-prone expressions are not rule-governed is simply preposterous. It's preposterous because such classifications are, of course, genuine classifications, apt to be *correct* or *incorrect*. And it seems the merest platitude that correctness, or incorrectness, has to be a matter of fit, or failure of fit, between an actually delivered verdict and *what ought to be said*. There is therefore no alternative but to construe vague classification as *in some sense* subject to norm — and now, if the correctness of our verdicts in general is to be a matter of competence, rather than accident, there seems no option but to concede that we are in some way masters of these norms and follow them in our linguistic practice.

I grant there is indeed no option but to concede that. The point, however, is that such a concession may not amount to very much — in particular, it may not amount to anything which sets up the trilemma we confronted above. In order to illustrate how this may be so, we can invoke a comparison with some thoughts about truth which I have canvassed in other work.[14] To speak the truth is to 'tell it like it is', represent things as they are, state what corresponds to the facts. Understood in one way, these phrases are platitudes and incorporate no substantial metaphysics of truth, at variance with, say, coherentist or pragmatist conceptions. The patter of correspondence, platitudinously understood, should motivate no questions about the nature of *facts* — what kind of entities they might be (as it were, sentence-shaped objects?) or how they might somehow be fitted to *correspond* in an appropriate way to beliefs and thoughts. I want to suggest that, in a similar way, the conception of basic classificatory linguistic practice as consisting in learning rules and following them is likewise open to a minimalist, or platitudinous construal, but that the trilemma:

<semantic incompleteness, incoherentism, epistemicism>,

arises only on a richer, non-deflated construal. The proper understanding of the idea that such classificatory competence is unprincipled is exactly that the relevant kind of richer construal is inappropriate.

So: that's the *shape* of a proposal. The question is how to fill it out.

VI The Modus Ponens Model of Rule-Following

Here is one possible direction — one that takes us to the heart of the agenda of the *Philosophical Investigations*.

Take as a simple, uncontroversial example of a rule-governed practice the case of Castling in chess. The rule states (something like):

If the squares between the King and one of its Rooks are unoccupied, and if neither

---

[14] See for instance "Truth: A Traditional Debate Reviewed" in supplementary volume 24 (1998) of the *Canadian Journal of Philosophy* (special issue on Pragmatism, guest edited by Cheryl Misak) 31-74; [reprinted in *Truth*, ed. Simon Blackburn and Keith Simmons, Oxford: Oxford University Press 1999, 203-38

> the King nor the Rook have previously moved in the course of the game, and if the King is not in check, nor would move through or into check, then it may be moved two squares toward the Rook and the latter then placed on the adjacent square behind it.

Following this rule involves assuring oneself that the antecedent of the conditional articulating the requirement of the rule is satisfied in the circumstances of a particular game, and then —if one chooses — availing oneself of the permission incorporated in the consequent. Generalising, what is suggested is a model of rule-following which involves (implicit) *reasoning*, of the form of a *modus ponens*, from a conditional statement whose antecedent formulates the initial conditions of the operation of the rule, and whose consequent then articulates the mandate, permission or prohibition that the rule involves. We can call this the *modus ponens model* of rule-following.[15]

The qualification, 'implicit', is suggested because the appropriateness of the *modus ponens* model is not restricted to cases where rule-following is informed by *self-conscious* inference. Following a familiar rule may be, very often, phenomenologically immediate and unreflective. But the important consideration, as far as the appropriateness of the *modus ponens* model goes, concerns what the rule-follower would acknowledge as justifying his performance. It is enough, in order for the *modus ponens* model to be appropriate, that the explicitly inferential structure of reasons it calls for should surface in that context. A practised chess player may decide to castle without any conscious thought but that of protecting his King from an attack down the left flank. But if the legality of the move were questioned, he would be prepared to advert explicitly to the relevant pattern of reasons of the *modus ponens* type, whose conditional premise embodied a formulation of the rule for Castling.

Plausibly, the *modus ponens* model thus extends a long way into the class of phenomenologically immediate decisions and judgements. How far does it extend? In particular, can the competent application of an expression *always* be conceived as rule-following in accordance with this simple prototype? More specifically still, can we regard the classifications we effect using basic vague expressions, of the kind typified by the usual suspects, as an example of rule-governed activity in accordance with the *modus ponens* model?

The answer, I believe, is no, and the reasons why not are instructive. Suppose, to the contrary, that competent classification using "red", e.g., is a rule-governed practice in the sense articulated by the *modus ponens* model. Then there should be a conditional which

---

[15] This idea, and its limitations, are anticipated in section V of my "Wittgenstein's Rule-Following Considerations and the Central Project of Theoretical Linguistics" in *Reflections on Chomsky*, ed. Alexander George (Oxford: Blackwell 1989), 233-64; reprinted at 170-213 of C. Wright *Rails to Infinity* (Cambridge, Mass.; Harvard University Press 2001.)

specifies the conditions under which predication of “red” to an item should be assented to, and such that each competent classification can be seen as (implicitly) inferentially grounded in the recognition, of a presented item, that it fulfils the antecedent. The picture, in other words, is that each informed, competent classification of an item, x, as red is underwritten by a pair of reasons of the form:

> If something is X, then it should be classified as “red”

and

> x is X

So OK: what is X? What is the property whose instantiation underwrites the proper application of “red” as fulfillment of the antecedent for the rule for Castling underwrites the judgement that the particular situation in the game is one in which one is permitted to Castle? It doesn’t seem that we know of any plausible candidate answer, apt for all cases, except to identify X with *red*. The same goes for the general run of predicates that make up the usual suspects: basic, vague predicates used to record the results of casual observation. In such cases, the correct answer to the question, what is the condition common to the minor premise and the antecedent of the conditional for the *modus ponens*?, seems to be irreducibly homophonic.

This observation has a striking consequence. It is that the price of continuing adherence to the *modus ponens* model in these cases is that we are forced to think of grasp of the concept of what is to be, say, red as *underlying and informing* competent practice with the predicate 'red'. If the explanation of competent practice with “red” adverts to the intention to follow a rule grasp of which demands a prior understanding of what it is for something to be red, then the latter grasp has to stand independent of the linguistic competence. So in effect, we commit ourselves to the picture of language mastery displayed in the quotation from Augustine with which Wittgenstein begins the *Philosophical Investigations* — not, to be sure, in the aspect of that picture which involves thinking of the semantics of expressions generally on the model that of a name and its bearer — but rather the aspect which we find made explicit only later, at §32:

> And now, I think, we can say: Augustine describes a learning of human language as if a child came to a strange country and did not understand the language of the country; that is, as if he already had a language, only not this one. Or again: it is as if the child could already *think*, only not yet speak. And “think”would here mean something like “talk to himself”.

Wittgenstein, of course, intends his reader to take on board the thought that Augustine prototypically committed a major philosophical mistake. If we agree with him and simply

repudiate the Augustinian picture, then we will have to conclude that, at the level of the basic, casual-observational classifications expressed by the usual suspects, the *modus ponens* model is inappropriate. The classifications effected by means of vague expressions of this basic kind are *not supported by reasons*, for the only possible candidates to constitute their reasons would demand, when so conceived, that we think of grasp of the concepts expressed by the vague predicates in question as something prior to and independent of the mastery of these predicates.

In his 'later' work Wittgenstein gives expression to an epistemology of understanding wherein language mastery is systematically conceived not merely as a means for the expression of concepts but as the medium in which our possession of them has its very being. As a reaction against the diametrically opposed view, that thought — at whatever degree of sophistication — may always intelligibly be conceived as constitutively independent of the thinker's possessing means for its expression, Wittgenstein's stance is surely compelling. But neither polar view—that language is merely the means of expression of thought, and that nothing worth regarding as thinking is possible without language, respectively—is correct. The correct view of the matter overall is presumably something more nuanced. There is much possible intelligent activity which, though wordless, would naturally call for explanation by ascription to the agent involved of conceptual — content-bearing — states of varying degrees of sophistication. Imagine, for instance, a chimpanzee who, off its own bat and after many days of manipulations in the corner of his cage, is suddenly able repeatedly to solve Rubik's cube; or the wonderful behaviour, even in an unfamiliar city, of a well-trained guide dog. On the other hand there are plenty of concepts — take, for instance, that of a parametric variable in a step of existential elimination — of which it strains credibility beyond breaking point to suppose that they could be fully grasped by someone with no linguistic competence at all. So if its implication of the Augustinian picture when applied to "red", "bald", "heap", etc., is to be sufficient reason to reject the *modus ponens* model of competence in those cases, we need a special consideration why. Is there one?

Fully to vindicate the assumption, which I will from hereon make, that Wittgenstein is right at least about the concepts to which we give expression by basic vague expressions—those for which retention of the *modus ponens* model would demand siding with Augustine—would need detailed further discussion. Here let me just gesture at one line of thought. We can certainly imagine a creature without language — the chimpanzee again — exhibiting some sort of concepts of colour: he may, for example, manifest a preference, among variously coloured but identically smelling and tasting boiled sweets, for green ones. More generally, there are any number of kinds of colour-sorting behaviours that could be — and in many cases, are — exhibited by pre-linguistic children. But grasping the colour concepts *we actually have* is not merely a matter of dispositions of appropriate response to paradigms. To grasp any classificatory concept, one needs not just to learn to respond appropriately to central cases but also to acquire a sense of its *limits*. With the usual suspects, however, it is

the very pattern of our *linguistic* practices that sets the limits, imprecise though they may be. We learn by immersion in the language how far one can stretch from paradigm cases of red, or blue, before classification starts to become *acceptably* controversial or difficult. Pre-linguistic children no doubt have some sort of grasp of colour saliencies. But the raw concepts they have, or could have, do not match the vagueness of colour concepts as linguistically captured. The reason is that it is precisely the onset of hesitancy, disorder and weak conflict in the *linguistic* verdicts of the competent that constitutes the gradual intrusion into the borderline area and sets the limits of the concepts in question. I do not deny that the chimpanzee might behave in ways which went some way towards giving sense to the idea that the concepts he was working with were vague — he might hesitate over a turquoise sweet, for example, while vigorously discarding a blue one. What seems to make no sense, however, is that his concept of green so manifested might be identical, vagueness and all, to the concept which, miraculously endowed with the power of speech and perfect basic English, he went on to display in his uses of "green". Competence with vague concepts is an essentially linguistic competence because the extent of their vagueness is an essentially linguistically manifest, socially constrained phenomenon.

If the broad direction of these remarks is correct, we must abandon the *modus ponens* model if we are to understand the sense in which the exercise of vague classifications is a form of rule-governed activity. One reaction would be to deny that it is, properly speaking, rule-governed at all — that it is sensible to think of an activity as *governed* by a rule only if one could in principle articulate the content of what would be the appropriate rules in such a way that a suitable thinker could *generate* a competence by observing them, and justify his performance by their lights. If I'm right, that condition is indeed unattainable where competence with vague classifications is concerned. That competence is, precisely, *not* to be viewed as a product of any possible anterior body of information which, in principle, could be used explicitly to inform the moves that competence requires. But actually, I don't think it matters whether we say that there is, properly speaking, no rule-following in such cases or whether we say merely — as I originally announced — that we need a more minimalistic conception of what in the relevant cases rule-following involves. Certainly, even with judgements involving vague expressions, there are still all of correctness and incorrectness, criticisability, proper responsibility, the intention to get things right, and a wide range of contexts in which it is important to succeed in that intention. But what there is not is a body of information which underlies competence, in a way in which knowledge of the rule for Castling together with knowledge of the history of the particular game and the present configuration of pieces enables (what may well be an unreflective) awareness by an expert player that Castling is an option here if he wants.[16]

---

[16] Additional difficulties for the *modus ponens* model are raised in Paul Boghossian's "Meaning, Rules and Intention," forthcoming in *Philosophical Studies*.

If this is right, then we can see our way round the trilemma. The trilemma arises if but only if it is a good question concerning the rules which govern our competence with "red": what do they *really have to say* when our colour classifications fall into the complex patterns which show that we are in the borderline area? — what is their message, what would we do if we were to do exactly what they require and nothing else? That the three answers canvassed seem, between them, to exhaust all of the alternatives, yet each to be objectionable, might have made one suspect that there is something wrong with the question. Now we see how that suspicion might be substantiated. There *is* no requirement imposed by the rules — not if we understand such a requirement as something whose character may be belied by our practice in the borderline area and into which there is scope for independent inquiry. There is no such requirement because to suppose otherwise is implicitly to commit oneself to the *modus ponens* model of rule-following for the classifications in question, and hence, as we have seen, to open an — at least locally —unsustainable gap between conceptual competence and the linguistic capacities that manifest it.

VII The Sorites Paradox of Observationality

To abandon the *modus ponens* model of competence with vague expressions is to open the way for the thought that the characteristic manifestations of vagueness in linguistic practice are, in a sense, *the whole story*. Vagueness is constituted at the level of use, rather than at the level of explanation of use.

Here are some of the kinds of fact that are salient at the level of use. First, competence with a vague expression is, in the usual run of cases, mastery of a practice which is communicated not by verbal explanations and characterisations but by immersion in that very practice. A large part of the acquisition of competence with a vague expression will typically consist in learning to use it to make judgements which we think of as immediately responsive just to how things strike us. Such judgements impress as, in a sense, undiscussable; we have to hand no repertoire of reasons with which to negotiate about them. Second, there is only such a thing as competence with such expressions because the relevant kind of training does generate — fortunately — a high degree of intersubjective — and cross-temporal intra-subjective — constancy in judgements, both positive and negative. Third, however, the training we receive also results in ranges of cases where constancy breaks down, where otherwise perfectly competent subjects differ, where views are characteristically unstable, and where we often find it difficult to be moved to a view at all.

My suggestion is that a philosophical account of the nature of vagueness is to be sought in the, by all means properly refined and nuanced, elaboration of facts such as these. There is no scope for an, as it were, information-theoretic standpoint whose goal is to outline rules which we implicitly follow in our use of vague expressions, and thereby to explain what

it is about those rules that gives rise to the characteristic manifestations. There is no such legitimate explanatory project. Competence with a vague expression is not the internalisation of a set of rules which prescribe the agreed verdicts in cases where agreement occurs, and whose other features explain the cases where it breaks down — in any of the ways that incompleteness, epistemicism and incoherentism respectively offer. The breaking down of constancy has to do, not with incompleteness, or inconsistency in the rules we tacitly follow, nor with the invisibility to us of the lines they draw, but is a matter of how we naturally respond, without reasons, to cases of significant distance from paradigms.

This shift in perspective provides for a major change in the impression given by Dummett's observational Sorites. Take what is surely the most daunting type of example: the Sorites associated with predicates like "looks red" over a series of pairwise indistinguishable square colour patches, running from, say, crimson to orange. Here adjacent patches *look just the same*. And "looks red' is surely an observational predicate in Dummett's sense if anything is. But observationality must require — must it not? — that the meaning of "looks red" enjoins that it is properly applied to both, if to either, of any pair of things which look just the same. So the paradox seems iron-cast. Yet this appearance precisely depends on a tacit construal of the observationality of "looks red' and its kin which rests on the *modus ponens* model of competence in their use. In effect, we are seduced into thinking of competence as involving the internalisation of a set of rules which connect appearances with the proper application of the predicate: rules of the form, roughly,

If x appears thus-and-so, then it should be classified as "looks red".

Once into this way of thinking of the matter, it is indeed impossible to understand how things which appear the same could possibly deserve any but the same classification. But the allure of the major premise dies away once the idea is taken on board that the classifications effected by these predicates are, in basic cases, essentially unsupported by any articulatable structure of reasons — more specifically, that they can be and normally are competently made without the mandate of any internalised conditional rule of the schematised kind. Since they need no such mandate, they are not properly regarded as unprincipled — at least not just on that account — when they collectively take a shape over a Sorites series — as they must if they are to be consistent — which would violate such a rule. Competent classification over a Sorites series for "looks red" will — must — at some point involve differential responses to presentations that look exactly alike. The (crazy) idea that competence somehow accordingly involves *disrespecting* the rules is an artifact of a misplaced adherence to the *modus ponens* model at a level where it involves an incoherent over-rationalisation of our practices. As Wittgenstein (near enough) says, not everything judged rationally is judged for reasons.[17]

I said earlier that the natural reaction to the Dummettian Sorites is to seek "an improved understanding of what an expression's possessing observational content involves". I

[17] *Investigations* §289

meant that reaction to contrast with one of mistrust of the very notion of an observational predicate. To be sure, if what I have been saying has any truth in it, then we should certainly mistrust the idea of an expression's being governed by rules which prescribe like verdicts about like appearances. But that is no cause for suspicion of Dummett's own working characterisation of observational expressions as those: "whose application can be decided merely by the employment of our sense organs".[18]. What is wrong is the little piece of theory which, under the aegis of the *modus ponens* model, links the Dummettian notion with governance by rules prescribing like verdicts about things our sense organs cannot distinguish. But the right reaction is not to retain the *modus ponens* model and look to somehow complicate or refine the rules which are characteristic of the competent employment of an observational predicate. It is to realise the implications of the point that competence with an *observational* predicate is not, even tacitly or "in principle", an *inferentially-controlled* competence.[19]

Dummett put the spotlight on observational predicates and I too have spoken of the usual suspects as predicates of 'casual observation'. But if the general drift of the foregoing is correct, vagueness will be an expectable characteristic of a more inclusive class of predicates (or other kinds of expression): all those which give rise to judgements for which competence is not, in basic cases, to be explained in terms of the *modus ponens* model — i.e., judgements competence with which is not a matter of sensitivity to inferentially-organised reasons but which are characteristically both rationally and phenomenologically immediate. Observational judgements are a species, but only one species, within this wider genus. My conjecture is that all vagueness of the kind which interests us here — vagueness of the kind which seems to be associated with higher-order vagueness and seems to give rise to Sorites-susceptibility —originates at the level of non-inferential judgement.

VIII Denouément

Unfortunately these considerations do not yet take us out of the wood. They provide the means to address one *motivation* for the major premisses in Sorites paradoxes for observational predicates. If they are compelling, and if that — illicitly rationalistic — motive exhausts the field, then we can indeed explain away the plausibility of those premises— an

---

[18] Cf. footnote 3

[19] The point I am making has I believe close connections with what Wittgenstein is driving at when he speaks in the famous passage at *Investigations* ß201 of "a way of grasping a rule which is *not* an *interpretation*, but which is exhibited in what we call "obeying the rule" and "going against " in actual cases." It would not stray very far from the intent of that passage, in my opinion, if we gloss it as: there is, and has to be, a kind of rule following in which the ingredient steps are performed without the possibility of an articulated justification in the light of a statement of the rule—, and in that sense are performed *blindly*. (Cf. *Investigations*ß219.)

essential part of any solution to a paradox, properly so regarded — and are now free, presumably, to regard each such Sorites simply as a *reductio* of its major premise. But now a new concern surfaces. To deny the major premise in the "looks red" Sorites, for example, is to conclude that it is *not true* that each patch that looks red in the series we imagined is succeeded by a patch that looks red..... But *how* is that not true? After all, we precisely so conceived the example that no patch that looks red is succeeded by one that *does not*. Surely no implicit over-rationalisation of competence with observational predicates is involved in the thought that no pair of indistinguishable patches can be such that one looks red while the other does not....?

It would be relevant to observe that treating the reasoning of the paradox as a *reductio* of a major premise of the form:

$(\forall x)(Fx \rightarrow Fx')$

is a commitment to affirming the corresponding claim of the form

$(\exists x)(Fx \ \& \sim Fx')$

only on the assumption that classical logic is good for vague statements. And for what it is worth, I do maintain that something like an intuitionistic logic — at the least, a logic in which double negation elimination does not hold unrestrictedly— will be required in any fully satisfactory treatment (and will say more on this shortly.) But someone could agree about that and still have the misgiving just aired. Even if we have disposed of the motivation for the universally quantified form of major premise in the observational Sorites, there still appears to be every motivation for the negative existential major premise for a corresponding observational instance of the No Sharp Boundaries paradox. Again: surely there just *nowhere occurs* in the relevant kind of series any patch that looks red while its immediate successor does not. Is not the negative existential major premise just flat-out true in this case?

The problem is, indeed, more general. Nothing has so far been said to address the overarching motivation, noted at the conclusion of section I, for the major premises in No Sharp Boundaries paradoxes as a class — namely, the simple thought that vagueness *per se* consists in absence of a sharp cut-off in the relevant series, hence in the truth of a statement of the form,

$\sim(\exists x)(Fx \ \& \sim Fx')$

What *is* vagueness if it is not a truth-conferrer on statements of that form?

These are very awkward looking questions. But I think the work we have done has taken us closer to being able to answer them. Let me conclude by sketching how.

In dropping the *modus ponens* model, we drop the idea that truth for vague statements is a matter of fit with the requirements of semantic rules which competence internalises. Rather, truth must now be viewed as grounded in the patterns of linguistic practice which competence, uninformed by any such rules and reconceived purely as the ability to

participate successfully in that very practice, exhibits. There is a temptation to try to say something detailed and specific about the nature of this grounding — perhaps in the form of a developed account of some form of response-dependence. I do not know if any such attempt can be fully successful. But we can get some mileage out of a theoretically more modest standpoint. Vagueness, of the kind that interests us, gives rise to definite cases as well as borderline ones. There are things that definitely look red, and things that definitely don't. These definite cases are cases which elicit a firm and stable consensus in judgement from those considered competent in conditions considered as relevantly suitable. And in discarding the *modus ponens* model, we now treat that sociological fact as, in a sense, the whole truth about the cases in question: there is nothing deeper, in particular no tacitly understood set of guiding requirements which explain and underwrite the convergence in verdicts about them, and no sense accordingly in the idea that the convergence might somehow be defeasible by further standards, implicit in the meanings of the expressions in question. The way is therefore open to us to regard such convergence as *unimprovable evidence* of the truth of the statements in question, and —*pari passu* —inconsistency with verdicts which enjoy such convergence should be regarded as unimprovable evidence of falsity.

The immediate effect of that consideration is that there is no option but to regard the verdicts about the end-points of a Sorites series — schematically, F0 and ~Fk — as straightforwardly true. For there is now nothing on which their truth might depend which is left out of the picture after competent convergence on them in good enough conditions is factored in. But that entails that there really is no alternative but to regard the negative existential major premise

$\sim(\exists x)(Fx \,\&\, \sim Fx')$

as false. It must be regarded as false because it is inconsistent with a consistent pair of statements for whose truth there is unimprovable evidence. Correspondingly, we have therefore no option but to accept *its* negation

$\sim \sim(\exists x)(Fx \,\&\, \sim Fx')$.

This realisation sets us two problems. The first is to understand how we *can* deny the negative existential major premise without controverting the fact of the vagueness of F within the relevant series. The second, closely related, is to explain how, after that denial, we can now avoid an apparently preposterous — *pace* Williamson —endorsement of the Unpalatable Existential

$(\exists x)(Fx \,\&\, \sim Fx')$,

which seems tantamount to an endorsement of F's precision.

First, then, on denying the negative existential.. Doesn't the vagueness of F in an appropriate series just consist in the absence of a sharp cut-off and isn't that just what the negative existential directly states? Well, no — or rather 'yes' to the second question and 'no' to the first. Vagueness, we have to learn, does *not* consist in the absence of sharp cut-offs. What the negative existential states — even in intuitionist logic — is the same as what is

stated by

$(\forall x)(\sim Fx' \multimap \sim Fx)$

which, in the presence of ~Fk, is tantamount to the false statement that there are no Fs in the series. So that *cannot* be the correct way to characterise the vagueness of the boundary between the Fs and the non-Fs in a series which contains both. What *is* a fact is the absence of any pattern of responses that would provide evidence for the *existence* of a sharp cut-off. The relevant pattern of responses would involve convergence on the initial elements of the series up to some last positive verdict, Fi, followed by convergence on the negative verdict, ~Fi', and then sustained convergence on negative verdicts up to the end. What we have instead is just the pattern of judgements distinctive of vagueness: polar constancy flanking the gradual breakdown of convergence in verdicts of any kind in the borderline region. That indeed, on the present perspective, is what vagueness consists in. So it doesn't consist in anything which makes the negative existential true. Rather, if the existence of a sharp cut-off is identified with the truth of the unpalatable existential, then we should say this: that what is distinctive of a predicate's vagueness is not that there is no sharp cut-off, but rather that there is nothing in our practice with the predicate that grounds the claim that there is a sharp cut-off (and that —again, *pace* Williamson —we have no satisfactory conception of how that claim might be grounded in an independent theoretical way.)

So as far as the first problem — that of explaining how we can deny the negative existential without controverting F's vagueness — is concerned, then, I am saying the following: that, so far from the negative existential's being a satisfactory characteristic statement of a predicate's vagueness, our practice with vague expressions provides unimprovable evidence that the negative existential is false, and hence that its negation — the double negation of the statement of the existence of a sharp cut-off — is true. The pattern of practice constitutive of vagueness, so far from providing evidence for the truth of the negative existential, is inconsistent with it and needs to be described quite differently.

As to the second problem, that of avoiding endorsement of the Unpalatable Existential,

$(\exists x)(Fx \ \& \sim Fx')$

we have so far observed merely that nothing in our patterns of judgement about each Fi provides any evidence for it, and that — *pace* Williamson — we have no other theoretical reason for thinking it true. These, on the present, "post- *modus ponens* model" perspective, are constitutive facts: vague expressions are ones for which sharp boundaries are neither drawn by actual competent practice nor posited by any known well-motivated theory. So there is no reason whatever to assert that there are such boundaries.[20] The temptation to think

---

[20] This claim presupposes, of course, that classical logic does not provide well-motivated theoretical reason to affirm the Unpalatable Existential on the ground of our acceptance, just argued for, of its double negation. That is exactly my position.I shall say more about it in a moment.

that this adds up to enough to justify *denying* the Unpalatable Existential—and so asserting the negative existential major premise—is very strong. It is of a species with a general tendency to deny things for which we have found absolutely no evidence (— the spirit in which we would, many of us, deny that there are leprechauns.) Usually, these somewhat arrogant denials do no serious harm. In this context, however, their nemesis is the No Sharp Boundaries paradox.

Let me relate these proposals to specific case of the No Sharp Boundaries paradox for "looks red". We left discussion stalled on the thought that it seems merely to be the simple truth that no patch in the relevant series which looks red is followed by one which does not. But that is not the simple truth. What *is* simply true is that nobody competent will willingly make a judgement which identifies such a pair of patches. And since none of us will do that singly, there is no question of a convergence on a cut-off. So nothing in our practice provides a ground for the claim that there is a cut-off, nor do we have any defensible conception of how there might be one nonetheless which our practice fails to reflect. On the other hand, the equivalent of the "simple truth", namely that no patch in the series which does not look red is immediately preceded by one which does, entails, when some of the later elements do in fact not look red, that none of them do. So the "simple truth that no patch in the series which looks red is followed by one which does not" is actually inconsistent with the way the elements collectively look and is therefore not a truth of any kind. As before, what has happened is a confusion of the complete absence of grounds for the existence of a cut-off with the possession of grounds for denying one.

But in this case I think there is an additional seductive fallacy. It is that we confuse its *not looking as if there is a sharp cut-off* in the series — which is true: if it did look as if there is a cut-off, we'd converge on it — with it's *looking as if there is none*. The latter is false. It does *not* look as if there is no cut-off — at least not if that is taken to mean that the series collectively looks as it would if there was indeed no cut-off. For if there were no cut-off, the series would have to look to be uniformly composed of squares which all looked red, or all did not look red. Which by hypothesis it does not.[21]

We are having to draw some quite subtle distinctions. It may help to consolidate a sense of what I am proposing in general if the reader reflects that, assuming 'monotonicity' for the series in question, the Unpalatable Existential is equivalent to the principle of

---

[21] The conflation is comparable to one at work in the idea that one can get an inkling of what experience of disembodied survival of death would be like by suffering total sensory and proprioceptive anaesthesia. The latter might indeed generate experience in which one's body was in no way presented to one. But its not appearing as if one has a body is not the same thing as a pattern of experience which represents one has having none.

Bivalence when restricted in range to each ingredient statement, Fi.[22] The temptation to deny the Unpalatable Existential is thus on all fours with the temptation to think that vagueness should prompt denial of Bivalence. That denial however — at least when truth is disquotational — likewise familiarly leads to contradiction. I am suggesting an attitude to the Unpalatable Existential which I take to be comparable in all essentials to the Intuitionists' attitude to Bivalence in those areas of mathematics where they regard it as unacceptable. It is there regarded as unacceptable because, when the truth of statements to which Bivalence is to be applied is conceived as constrained in ways that are taken to be independently motivated, it *goes beyond all evidence* to suppose that each instance of Bivalence is true. In the case of mathematical statements, the Intuitionists take it that there is independent motive to require truths to be constructively provable. I do not know whether the historical motivation for that thought is wholly independent of what I have proposed about vague statements: that their truth and falsity have to be thought of as determined by our very practice, rather than by principles which notionally underlie it. But the effect is similar. The denial of any instance of Bivalence leads by unexceptionable basic logic (and the usual truth-rules) to contradiction. So we must deny any such denial. But when Fi is borderline, the break down in convergence of verdicts leaves us without uncontroverted evidence for its truth or for the truth of its negation. Since we lack convincing theoretical grounds to think that one or the other must be true nonetheless — because such grounds would have to regard the determinants of truth and falsity as constituted elsewhere than in our linguistic practice — we are left with no compelling reason to regard either Fi or its negation as true. We should therefore abstain from unrestricted use of the law of double negation elimination. We do so on the grounds that it is non-conservative of knowledge — exactly the Intuitionistic reservation.

The thought that vagueness might require a non-classical logic is, of course, hardly original. But it has usually been clothed in broadly semantic proposals about third truth-values, truth-value gaps and degrees of truth. All these proposals see the principle of Bivalence as breaking down where vague statements are concerned. I am suggesting that the correct basic complaint to have about Bivalence where vague statements are concerned is not that it breaks down but that it goes beyond the evidence. I have supported this contention by argument that vagueness is a phenomenon of judgement unsupported by reasons, and that this makes for an especially intimate connection between patterns in competent linguistic practice and truth. It is this intimacy of connection that allows us to rest content with our stable verdicts about the poles of a Sorites series and so to enforce denial of the major premises involved. It is the same intimacy of connection that allows us to say that there is no reason — not merely: no reason provided by our actual patterns of judgement, but no reason whatever

---

[22] For if $(\exists x)(Fx \ \& \ \sim Fx')$ is true, then monotonicity ensures that each $y < x$ will be F and each $z > x'$ will be non-F. So every element will be determinately either F or non-F. Conversely, if Bivalence holds, and there are both Fs and non-Fs in the series, then therewill have to be an adjacent pair, the earlier F and the other not.

— to regard the Unpalatable Existential as true. So double negation elimination is, in this instance, epistemically non-conservative.

There is a lot more that would need to be discussed and developed in a fully explicit treatment but I must draw this particular essay to a close. It is, of course, conspicuous among Michael Dummett's many philosophical achievements to have done more than anyone else to rehabilitate the Intuitionistic outlook and to transform it into a significant contemporary force not just in the philosophy of mathematics and logic but in metaphysics and the theory of meaning generally. If there is anything in what I have been saying, he may, ironically, have overlooked one especially apt and helpful application of it, to the riddle of vagueness.

Arché, AHRC Centre for the Philosophy of Logic, Language, Mathematics and Mind, University of St. Andrews,
and New York University