CHRISTOPHER PEACOCKE

ARE VAGUE PREDICATES INCOHERENT?*

Does the Sorites paradox show a wide class of observational expressions to be incoherent? Michael Dummett has argued that it does.[1]

To accept Dummett's argument is not to be forced to abandon observational predicates altogether. His argument does not apply to all observational predicates: one can retain 'is discriminably different from' and such comparatives as 'is yellower than' and 'is balder than'. But clearly we could not express everything we originally wanted to express. A description of an object using vocabulary to which Dummett's argument does not apply will not always settle the question of whether it is green, or whether it is bald if it is a person. My aim in this paper is to offer a diagnosis which does not blame the Sorites paradox on the incoherence of certain vague predicates, and which allows it to be literally true that an object is green.

First I will consider some other reactions to Dummett's argument. They are reactions with which it is hard to rest content. In considering why this is so we will discover properties which must be possessed by any more satisfying reaction.

I

Crispin Wright has suggested that one possible reaction to Dummett's argument is to say that the paradox establishes the following conditional conclusion: if we regard understanding an expression as grasping certain kinds of rules which are to govern its use, then these vague observational predicates are incoherent. This theorist Wright envisages contraposes and concludes that such a conception of understanding is mistaken. The use of such predicates cannot be completely determined by a set of incoherent rules; for, the theorist will say, "our use of these predicates is largely *successful*; the expectations which we form on the basis of others' ascriptions of

* I am greatly indebted to Crispin Wright for comments on an earlier draft of this material.

colour are not normally disappointed. Agreement is generally possible about how colours are to be described . . . ".[2] So he says that "the methodological approach to *these* [vague observational] expressions, at any rate, must be more purely behaviouristic and anti-reflective, if a general theory of meaning is to be possible at all" (ibid., p. 247).

This suggestion is not sufficient to defuse the Sorites paradox. For consider this predicate C of objects: $C(x)$ iff x is such that the community will agree in calling it 'red'. Now suppose too that a difference d in the wavelength (w/1) of light is not visually discriminable by any member of the community; and that light of w/1 k is definitely red. Then we can still construct this paradox:

> If a reflects light of w/1 k, then $C(a)$.
> If an object differs in the w/1 of light which it reflects by just d from something that is C, it too is C.
>
> ---
>
> All visible objects (reflecting pure light) are C.

Here an absurd conclusion has been drawn in terms of the vocabulary *we* use for describing the linguistic practices of the community. The presence or absence of behaviourism is not obviously the problem: this paradox must be resolved even if the predicate C has a behavioural definition. Thus the paradox seems to arise even if we do not suppose that the use of these expressions is governed by rules.[3]

It might be objected that since 'C' is not an observational predicate, the second premise of this new paradox is not true; and so it might be concluded that the reaction Wright considers is not vulnerable to this difficulty. Now it is certainly true that 'C' is not literally an observational predicate. But it is related to an observational predicate, viz., 'red', in such a way that the *reasons* given for saying that the second premise of a Sorites argument must be true are applicable to the second premise of this new argument too. The reason given for accepting the second premise of a standard Sorites argument using the predicate 'red' is that it would be inconsistent with the observationality of 'red' to suppose that a difference d of wavelength which is not visually discriminable could make the difference between a situation in which the predicate applies and one in which it does not. But it would be equally inconsistent with the observationality of 'red' to suppose that a difference d of wavelength, something *ex hypothesi* not visually detectable by any member of that community, can make the difference between an object being such that the community will agree in calling it 'red' and an object without that property.

One proposal for avoiding this metalinguistic paradox might be based upon a comparison between a person or a community using an observational predicate with a measuring instrument the readings of which are displayed in digital form. A given small change in the magnitude presented to the instrument may or may not produce an alteration in its reading, depending upon the internal state of the instrument: there will be an alteration if the instrument is internally sufficiently near a threshold of sensitivity. For such an instrument, an analogue of the major premise of the metalinguistic paradox would be false: whether the instrument produces a different response to a given change in what is presented to it must depend on its internal states. Can we not give a similar description of the linguistic practices of a person or community using a vague predicate? There is, however, a disanalogy between the cases which prevents the instrument from being used as a model to avoid the metalinguistic paradox.

An observational predicate is one whose application to an object can be determined from the kind of experience produced by that object in standard conditions: in particular, if two objects produce experiences which are not in quality discriminably different from one another, it cannot be that an observational predicate definitely applies to one of the objects and does not definitely apply to the other. Now any model for an application of vague observational predicates must provide analogues of three things involved in such application: there must be states which are the analogues of having experiences, there must be something analogous to the continuity of experience as reflected in the nontransitivity of nondiscriminable difference, and there must be some analogue of the application of an observational predicate upon a particular occasion. It seems impossible to provide for all three of these in the model of the digital instrument. Suppose we took the instrument's internal states as analogous to the having of experiences and the display of a digital reading as analogous to the application of an observational predicate. What could be the analogue of continuity of experience? Are we to say that the analogues of experiences which are not discriminably different in quality are internal states of the instrument produced by physical magnitudes differing by less than a specified amount *d*? This would destroy the analogy. There are internal states either side of the threshold corresponding to a particular reading which are produced by physical magnitudes differing by less than *d*. So under this analogy there would be no difficulty in the idea of an observational predicate (corresponding to one reading) definitely applying to one but not to the other of two objects which produce experiences which are not discriminably different in quality. But in fact we can make no sense of this idea.

It may seem tendentious to use the concept of experience in this argument against one attempt to avoid the metalinguistic paradox. Was not Wright explicitly concerned with a more 'behaviouristic' characterization of the use of language? But 'behaviouristic' here did not mean: behaviourally specifiable. It meant: not based on the suggestion that a speaker's application of predicates is governed by rules he uses to guide him in the use of language.

Suppose, perhaps *per impossible*, there were some way of blocking the paradox involving '*C*' which could not equally be applied directly or indirectly to the original object language paradox involving 'red'. Should we then be satisfied with saying: "The community will agree in calling certain things 'red' and not others. We can describe this state of affairs without paradox, but we cannot say under what conditions they will agree in calling something red."? No: to stop here would be to fail to say what information is conveyed by an utterance containing 'red'. A theory of meaning which did stop here would not be something which, if known, would put someone in a position to understand the language of the community.

Mark Platts has also reacted to Dummett's argument.[4] Platt says that "We grasp the use of a vague predicate at least in part through a group of paradigm exemplars of them."[5] If a patch of colour is not discriminable from a red patch, it is itself red. But, Platts holds, one cannot construct a Sorites paradox with the predicate 'is a paradigm red patch': he writes that "indiscriminability from a paradigm of red does not of necessity mean that we have *another* paradigm of red."[6]

This suggestion seems to ignore the role of observationality in the claims of Dummett and Wright. Dummett and Wright made a case that an observational predicate must be applicable to both or neither of a pair of objects which are not discriminably different. Having in no way undermined or qualified the principle, Platts' view is open to a simple dilemma. Either 'is a paradigm of red' is an observational predicate, or it is not. If it is, then the Sorites paradox can be restated with respect to it. If it is not, the account he gives of what it is to understand 'red' makes 'red' not an observational predicate, which seems wrong. Moreover, while the principle relating observationality and indiscriminability stands in unqualified form, the original paradox using 'red' itself has not been defused.[7]

II

It is sometimes suggested that the Sorites paradox can be neutralized by making proper use of the point that vague predicates are predicates of degree: it can be that one thing is red to a greater degree than another. Wright has argued that such considerations cannot block the paradoxes.[8] Let us initially explain the notion of degree thus: two objects are red (say) to the same degree iff any object not discriminably different from one of them in respect of colour is not discriminably different in respect of colour from the other. There are many ambiguities and indeterminanacies in this definition, but let us ignore them just at present, since the main point we want to make holds under all ways of resolving them. The important point is that the degree to which an observational predicate applies to an object is not itself an observational matter, in the following sense: two objects can be not discriminably different from each other, and yet the degree to which a given observational predicate applies to the two objects may be different.

There is a familiar argument for that conclusion. Suppose for *reductio* that it were not so; that is, suppose that any observational predicate applies in exactly the same degree to any pair of objects which are not discriminably different. Now consider a triad of red objects *a*, *b*, *c* where

> *a* is not discriminably different ("d.d.") from *b*
> *b* is not d.d. from *c*
> *a* is d.d. from *c*; in particular *a* is redder than *c*.

Then the degree to which 'red' applies to *a* must be the same as the degree to which it applies to *b*, under the hypothesis of the *reductio*; since the degree to which 'red' applies to *b* is similarly the same as the degree to which it applies to *c*, it follows that the degree to which 'red' applies to *a* is the same as the degree to which it applies to *c*. But that is not consistent with the fact that *a* is redder than *c*. Hence we must conclude that a predicate can apply in different degrees to objects not d.d. from one another.

The notion of degree we are using here is close to Goodman's concept of identity for qualia. Suppose we are prepared to quantify over colours as universals and to treat matching as a relation between such universals. Then we can say that object x and object y are red to the same degree iff any colour matching the colour of either one of them matches the colour of the other. Similarly, the degree to which x is red is greater than the degree to which y is red iff some colour matching the colour of x is redder than the colour of y. If we want to be free of the quantification over colour universals, or want at least to define the matching relation between them in terms of relations between particulars, there are severe difficulties if we do not want to take identity of shade as a primitive notion; but since these difficulties equally affect sharp observational notions, and so cannot be the source of the Sorites paradox, I will relegate them to a footnote.[9] In any case, we should note that although difference of degree to which an observational predicate applies is not always an observational notion, the notion of degree has been *explained* in terms of observational notions such as 'matching' (nondiscriminable difference), plus logical notions. In this sense the notion of degree does not go beyond distinctions manifested in the abilities exercised by the speakers of the language; this is in contrast with those who are prepared to employ, for instance, a sharp notion of a family of admissible valuations in giving a semantics for vague expressions.

If we relativise the major premise of the traditional Sorites argument by using this notion of degree, do we still obtain the paradox? When we relativise we obtained this:

> If an object is red and a second object is not d.d. from the first, then the second is red to a degree which one cannot determine just by looking at those two objects to be different from the degree to which the first is red.

Our point was that nevertheless these degrees may *be* different. So we cannot relativize the premises of the Sorites argument and show that an orange is red to the same degree as a British pillar box.

What of the original major premise in the colour version of the paradox? Is it still true that anything not d.d. from a red object is itself red? We now have to hand the materials for arguing that it can be simultaneously true that *a* is not d.d. from *b* while the conditional

if *a* is red, *b* is red

has a consequent with a lower degree of truth than its antecedent. I claim that the Sorites paradox shows that there cannot be a conditional possessing both of these properties:

(1) *modus ponens* inferences for this conditional are valid without restriction

(2) for some observational predicate *F*, and any objects *x* and *y* (named by *a* and *b* respectively), if *x* is not discriminably different from *y*, then the conditional with antecedent '*a* is *F*' and consequent '*b* is *F*' is true.

One can interpret the conditional in such a way that the major or conditional premises of the Sorites argument are true, by (say) counting a conditional as true if its antecedent and consequent do not discriminably differ in their degree of truth; but then of course *modus ponens* will not be unrestrictedly valid. Alternatively one can retain *modus ponens* and require (consequentially) that such a conditional is true only if its antecedent and consequent have exactly the same degree of truth: and then the major premise of a Sorites paradox is false. But one must stick to one of these two courses consistently. So my suggestion is that the paradox results from the use of a conditional taken to satisfy incompatible conditions, rather than from any incoherence in vague predicates. (In fact (2) here is stronger than is necessary for the properties to be incompatible: it suffices to use as antecedent the condition that *x* and *y* do not differ sufficiently for *F* to be applicable to one and not to the other of the two.)

The question now remains of how on this proposal we can give some positive characterization of observationality. It was, after all, such characterizations which led Dummett and Wright to find the major premise of the paradox so compelling. But there are ways of stating the connection between nondiscriminable difference and observationality which do not lead to paradox. In particular, we can say that if *F* is an observational property, then if two objects *x* and *y* are not discriminably different (in respect of a given kind of property), then it is not the case that *x* is definitely *F* and *y* is not definitely *F*; in symbols,

$$Ixy \rightarrow \sim(DFx \ \& \sim DFy)$$

This last conditional has a sharp antecedent and a consequent whose truth can be a matter of degree: it resembles "If the ball is in this urn, it is red". If F is an observational property, then whenever the antecedent 'Ixy' is true, the consequent will be true enough for the whole conditional to be true (all this relative to a given assignment to the variables). But to obtain a Sorites paradox from this conditional, we would need something equivalent to the principle

$$\frac{\begin{array}{l}\sim(DFx\ \&\ \sim DFy)\\ DFx\end{array}}{DFy}$$

This principle should be rejected if Ixy is to be sufficient for $\sim(DFx\ \&\ \sim DFy)$: for if y is F to a lesser degree than x, then the conclusion will have a lesser degree of truth than the premise DFx. Again, we should not be surprised to obtain paradoxes from this principle.

There is an apparent problem for this general diagnosis over the metalinguistic paradox. The predicate C is not a predicate of degree. For any given object, either all members of the community will agree in calling it 'red', or not all members will call it 'red'. Here there is no room for talk of vagueness or matters of degree. C is a sharp predicate, and it does not make sense to say that it is true in a higher degree of some objects to which it applies than it is of others to which it applies. Does not the metalinguistic paradox show then that my diagnosis does not cover all examples of Sorites-like paradoxes?

I reply that it covers the metalinguistic paradox indirectly. Sometimes we have to state a theory about the extension of a sharp predicate by using a vague predicate. In particular is this true of the sharp predicate C and the vague predicate 'red'. We have to say that:

> Any object which is red the community will agree in calling 'red';
> any object which is not red the community will not agree in calling 'red'.

I suggest that the only reason that we feel tempted to accept the major premise of the metalinguistic paradox is that we employ reasoning using conditions with properties (1) and (2) and which contain 'red', and then go on to apply these two general principles to draw conclusions which contain the predicate C.

Without using the conditional which has to have incompatible properties, we have no reasons for believing the major premise of the metalinguistic paradox. Insofar as inductive evidence might support that major premise, it must, if enlarged by further investigation, eventually refute the premise. Since *C* is sharp we have no option but to interpret the conditional in the major premise as a classical material conditional (rather than anything whose semantics mentions degrees), and then some instance of this premise must be false if the principles displayed earlier in this paragraph are true. Since *C* itself is not an observational predicate, there is no pressure against this conclusion: all we used earlier in raising the metalinguistic paradox as a problem for one position is that it shared a property with the observational predicate 'red', viz. that if *x* is definitely red and *y* is not discriminably different from *x*, then it is not the case that *y* is not red and not the case that *y* is not *C*.

Let us return to the object language paradox. To say that (1) and (2) are incompatible properties is not to imply that there is no interpretation of the conditional on which *modus ponens* is valid without restriction and on which 'if *a* red *b* is red' is sometimes true where *a* and *b* name observationally indistinguishable objects. We do not need to dispute the validity of the inference

b is red
―――――――――――
if *a* is red, then *b* is red.

nor that of

b is red
―――――――――――
either *a* is not red or *b* is red
―――――――――――
if *a* is red, then *b* is red

To account for the validity of these inferences, we need only to construe the conditional ⌜if *A* then *B*⌝ as true when either *A* is definitely false or *B* is definitely true, when *A* and *B* have definite truth values. For such a partially defined conditional, *modus ponens* will always preserve definite truth; this conditional does not have property (2). Indeed, for such a conditional the truth value of the whole depends just on the truth values of its constituents, and not upon their degrees of truth. Using a distinction of Dummett's we may say that the *content* sense and the *ingredient* sense of *A* and *B* with respect to this conditional are identical: "... we must distinguish between knowing the meaning of a statement in the sense of grasping the content of an assertion of it, and in the sense of knowing the contribution it makes to determining the content of a complex statement in which it is a constituent: let us refer to the former as simply knowing the *content* of the statement, and to the latter as knowing its *ingredient sense*".[12] In the case of a conditional the truth value of which has to be specified as a function of the *degree* of truth of its constituents, ingredient sense and content sense would come apart.

We might attempt to extend the interpretation of the conditional for which the inferences displayed in the last paragraph are valid so that its truth value is determinate in other cases too. But there is not a new source of paradox here. If the extended specification has property (2), as it would if we said that ⌜if *A*, then *B*⌝ is to be definitely true when *A* and *B* differ in degree of truth indiscriminably, then *modus ponens* will no longer be valid in general.

It would be implausible to claim that the "if" of English determinately either abandons property (1) or abandons property (2). In practice we are prepared to reject false conclusions arrived at by Sorites-like reasoning, while not being prepared specifically to blame either the form of the inference or some conditional instance of the major premise. As a consequence, no conditional with a determinate formal semantics and possessing either property (1) or property (2) can claim to express precisely the meaning of the English conditional used in the presence of vague predicates.

The vague expressions we have considered so far here have been observational predicates. But an expression can be vague and feature in Sorites-like reasoning essentially while being neither a predicate nor in any natural sense observational. The quantifier 'many' is an example. We can often reach false conclusions from the true premises of the form:

> Many *F*'s are *G*
> For any object *x* which is *F* and *G*, if many *F*'s are *G* then many *F*'s are *G* and distinct from *x*.

For example, suppose the complete list of members of some society is *a*, . . . , *z*, and that many members of the society, *c*, . . . , *z*, say, voted for the resolution. Repeated application of the two displayed premises can lead us from these true suppositions to the false conclusion that many members of the society voted for the resolution and are distinct from each of $c, \ldots, z$. The diagnosis of the paradox as resulting from the incompatible conditions placed on the conditional still applies, since whether many F's are G is a matter of degree.

It is most important that to argue for the coherence of some observational predicates where nondiscriminable difference is non-transitive is not thereby to argue for the coherence of a notion of an observational *shade*. Indeed, Dummett does show that the notion of an observational shade is incoherent. Let us consider just the case of colour. The concept of an observational shade is intended to conform to these three principles:

(i) if x and y are discriminably different, they are not of the same observational shade

(ii) if x and y are not discriminably different, they are of the same observational shade.

(iii) an object has at most one observational shade (at a given point on its surface, in the case of colour).

It is obvious that these principles lead to contradiction in the case of a triad a, b, c where a is not discriminably different from b and b is not discriminably different from c. a and c (at given points) have different observational shades (by (i)). But what shade can b have? It has to be both the same as that of a and the same as that of c (by (ii)); yet the shades of a and c are distinct.

The argument to this contradiction does not employ any conditionals whose consequents have a lower degree of truth than their antecedents. There is no obstacle to accepting this proof of incoherence while retaining our resolution of the Sorites paradox. (Ordinary colour predicates do not, of course, conform to a principle analogous to (i).) With this distinction between observational shades and vague predicates in mind, let us consider Dummett's example of the slowly moving pointer:

> I look at something which is moving, but moving too slowly for me to be able to see that it is moving. After one second, it still looks to me as though it is in the same position; similarly after three seconds. After four seconds, however, I can recognize that it has moved from where it was at the start, i.e. four seconds ago. At this time, however, it does not look to me as though it is in a different position from that it was in one, or even three, seconds before. Do I not contradict myself in the very attempt to

express how it looks to me? Suppose I give the name 'position *X*' to the position in which I first see it, and make an announcement every second. Then at the end of the first second, I must say, 'It still looks to me to be in position X'. And I must say the same at the end of the second and the third second. What am I to say at the end of the fourth second? It does not seem that I can say anything other than, 'It no longer looks to me to be in position *X*'; for position *X* was defined to be the position it was in when I first started looking at it, and, by hypothesis, at the end of four seconds it no longer looks to me to be in the same position as when I started looking. But, then, it seems that, from the fact that after three seconds I said, 'It still looks to me to be in a different position from that it was in after three seconds', that I am committed to the proposition, 'After four seconds it looks to me to be in a different position from that it was in after three seconds'. But this is precisely what I want to deny.[13]

What is going on here? Dummett introduces a notion of position which conforms to principles (i)–(iii); and then for that notion derives the contradiction as reached in the previous paragraph. He then remarks

One may be inclined to dismiss Frege's idea [that the use of vague expressions is fundamentally incoherent] if one does not reflect on examples such as these.

But this example points up only the incoherence of observational shades, and not vague expressions more generally. The argument Dummett gives in the case of observable position could not be reproduced for such predicates as "in the left of one's visual field" without using Sorites-type reasoning.[14]

III

Does it make sense to suppose that the world itself is vague? If so, does it undermine our resolution of the Sorites paradox?

It is natural to construe the suggestion that the world itself is vague as the suggestion that the world has to be described by (*inter alia*) vague expressions, where this need is not in some way a result of limitations on our capacities.[15] An uninteresting way of interpreting this suggestion is as the denial that vague expressions can have sharp translations. This is uninteresting because it is reasonable to require that translations preserve vagueness: so this interpretation of the suggestion hides substantive philosophical issues. One formulation which does not make the issue vanish is this. Suppose we have a language L containing vague expressions. Then the suggestion that the world itself is not vague is the suggestion that there will be some conceivable language L^1 which contains no vague expressions and which has the following property: it is *a priori* that if two situations agree in all respects describable using the language L^1, then they agree in all respects describable using the language L. This is a form of supervenience. I shall say that the vagueness of a vague expression *E* is *superficial* if for any language L whose sole vague expression is *E*, there is some language L^1 containing only sharp expressions, and such that the descriptions of L supervene on those of L^1 in the sense just explained. It would not be disputed that the vagueness of *some* expressions is superficial. The quantifier 'many' is an example. The truth values of sentences containing 'many' will supervene on those sentences not containing 'many' but containing cardinality quantifiers: there cannot, for example, be two situations with respect to one of which some sentence of the form 'Many *F*'s are *G*' is true and with respect to the other is false, if the situations have the same number of *F*'s being *G*. On this construal, then, the thesis that the world itself is vague would be the thesis that not all (possible) vague expressions have merely superficial vagueness.

I mention this interpretation to distinguish it, in thought at least, from another sense in which it might be said that the world itself is vague. This is the vagueness that is denied by the principle that for any simple property ϕ the presence of which in an object is a matter of degree, the relation "x is more ϕ (or ϕ-er) than y" is a total ordering: thus the degree to which one object is ϕ must be either greater, less or the same as the degree to which some other object is ϕ. There cannot be incomparable degrees. The restriction to simple properties would be difficult to make precise, but is clearly necessary for the principle to be plausible: no one would expect in advance the orderings 'x is a better novel than y' or 'x is a better city to live in that y' to be total, and this seems to be a result of the fact that so many different comparisons would have to be made before one could make reasonable judgements of such orderings between particular objects. There is no reason in advance why some combinations of these component comparisons should not result in there being pairs of cities which are incomparable in respect of which is better to live in, while one also has no reason to say that they are equally good. On the other hand, it may be said that it is one interpretation of the claim that the world itself is not vague that this situation cannot arise for simple properties. The claim is certainly plausible for a wide range of simple properties: we find it hard to make sense of the possibility that of two rods neither is physically longer than the other and yet they are not of equal length.

Are the degrees to which something may have the colour red totally ordered? If every coloured object had (at each point and at each time) a determinate Goodmanian shade, then, provided we hold such factors as brightness and saturation constant, there would be a total ordering of degrees of redness. But does every object have a determinate Goodmanian shade?

There seems to be a radical indeterminacy in applying Goodmanian shades to actual objects or experiences, of a kind which suggests that the circularity we discussed some pages back is indicative of an important point, rather than being a reflection of limited ingenuity in formulation. Suppose, to give the Goodmanian shades as favourable a chance as possible, we do not question that there are objects of any arbitrary Goodmanian shade (though this may already commit us to the existence of infinitely many things). Consider two objects *a* and *b* which are simultaneously perceived by the same person, and which are not juxtaposed and which are not discriminably different from one another. Suppose too that if they are moved into juxtaposition, there is a colour boundary between them – they are then discriminably different. (Alternatively for the conclusion of the present argument it suffices that some third object which matches one but not the other in colour is moved in between the two given objects and is in juxtaposition with both of them.) One might take the presence of this visible boundary as conclusive evidence that *a* and *b* had different Goodmanian shades all along, since things have identical Goodmanian shades only if they match in colour exactly the same things. But one would be wrong to do so. There is nothing in the situation as described to rule out the supposition that the Goodmanian shades of *a* and *b* were the same until they were moved into juxtaposition, and just before this moment the shade of one of them (which one?) altered indiscernibly so that they do not match when juxtaposed. There is a serious question whether there is anything for the difference between these two hypotheses to consist in. It is important that the problem here is a constitutive one, one of meaning, and not one of verification. Dummett has remarked "Although, as is well-known, some philosophers have gone down this path, it will seem quite unreasonable to deny that someone who was capable of telling, by looking or feeling, whether or not a stick is straight knew what it was for a stick to be straight, on the ground that he would not thereby show that he knew what it was for a stick which no one had seen or touched to be straight".[16] Straightness is a primary quality; if some object alters in respect of this property while it is unobserved, there is no difficulty in saying what the alteration consists in – it must involve some redistribution of matter in space. But constancy of secondary qualities must, in given observational conditions, depend on constancy of experience. An indiscriminable alteration in the shade of some object produces *ex hypothesi* no change in the experiences produced by the object. There seems to be nothing for such alterations to consist in. The impossibility this produces of assigning determinate, totally ordered Goodmanian shades to objects has nothing to do with the inexpressiveness or inadequacy of our language: the point remains however much we refine the language, for the point concerns the nature of experience itself.[17]

The conclusion to be drawn from these considerations is this, if we wish to continue to hold that all objects which are red are red to some degree or other, then we must confine ourselves to making those statements of degree which are true under *every* assignment of Goodmanian shades to objects consistent with the requirement that if one object is perceptibly redder than another, then it is red to a greater degree. But this has the consequence that in our example of *a* and *b* in the previous paragraph, neither is red to a greater degree than the other, nor are they equally red. For some assignments of Goodmanian shades (consistent with the requirement just mentioned) make *a* red to a greater degree than *b*, and some make *b* red to a greater degree than *a*. Thus the resulting final degrees of redness are not totally ordered.[18]

If this is correct, then our resolution of the Sorites paradox can stand: it is just that we must not naively take the degree to which objects may have some property as always totally ordered. One can accept this resolution while holding that on one interpretation of the phrase, it is indeed true that the world itself is vague.

New College, Oxford.

APPENDIX: FORMAL SEMANTICS

If 'red' were an incoherent predicate, we could hardly use it in any theory of ours, including a theory of truth for a language. Even if a theory containing it did not also have the resources for demonstrating the incoherence, there would be an acceptable extension of the theory that would have them. But if we accept that the previous considerations suffice to block the Sorites paradox, then there is no obvious objection to using 'red' in the metalanguage of a truth theory for a language containing 'red'. (It is not as if there were much else available which we could use in stating the semantic contribution of 'red' to the sentences in which it occurs, other than 'red' itself.)

Consider a fragment of a language with several proper names, the predicate 'red', and sentential negation. We want the degree to which the predicate "is true" applies to the sentence 'Uranus is red' to match precisely the degree to which Uranus *is* red. In the strictest form of representation, we would use here a variable-binding operator $\in^x_y$ which applies to a pair of predicates, which yields something which takes two terms to form a sentence:

$$\in^x_y[\phi(x), \psi(y)](t, t')$$

This formula would be true iff ϕ and ψ are true to exactly the same degree of the objects denoted by t and t'. (All this would need relativisation to a sequence in a first-order language.) Thus we want it to come out that

$$\in^x_y[\text{true}(x), \text{red}(y)](\ulcorner\text{red(Uranus)}\urcorner, \text{Uranus}).$$

We noted earlier that there is a qualitative aspect to the notion of degree on which vague predicates are predicates of degree. $\in$ is to be understood in such a way that the truth of this last formula requires the qualitative aspects of the degrees of the two predicates to coincide: if Uranus is not red enough to be said to be red, then $\ulcorner$red (Uranus)$\urcorner$ is not true enough to be said to be true (and so forth).

In fact in order to employ an operator of a syntactic category with which we are more familiar, I will write the last displayed formula with a special biconditional '$\leftrightarrow$' thus:

$$\text{true }(\ulcorner\text{red (Uranus)}\urcorner) \leftrightarrow \text{red (Uranus)}.$$

$\ulcorner A \leftrightarrow B\urcorner$ is true iff A and B are true to exactly the same degree. ('$\leftrightarrow$' thus has property (1) and lacks property (2) we discussed earlier.) Since we will here be concerned only with examples in which vagueness in sentences is produced by vague predicates, the substitutivity of identity will hold unrestricted in such contexts; indeed it is obvious given that we can regard the last displayed formula as an abbreviation of the one before it. '↔' is *degree-functional*: the degree to which ⌜$A \leftrightarrow B$⌝ is true is determined by the degree of truth of A and B. For closed sentences A and B. ⌜$A \leftrightarrow B$⌝ will be definitely true or definitely not true. But of course that will not be the case with other degree-functional connectives, for instance negation. This degree-functional negation we will write '⌝'. We shall not use any operators in the theory of truth which are not degree-functional. Degree-functionality is the analogue for a simple vague language of truth-functionality for simple classical languages.

We can then offer the obvious truth-theoretic axioms:

A1 $\forall t$(true (⌜red(t)⌝) ↔ red (den(t)))
A2 $\forall A$ (true (⌜⌝A⌝) ↔ ⌝ true (A))
A3 den (Uranus) = Uranus

We can derive the T-sentence for 'red(Uranus)' from A1 and A3, using the noted transparency of '↔'. By the degree functionality of '⌝' we can have inferential principles allowing the substitution of '↔'-equivalents within contexts governed by '⌝'. Hence from the already proved T-sentence for red (Uranus) and A2 we can derive that for ⌜⌝red(Uranus)⌝. Plainly there is nothing of much technical interest here.

What of the model theory? One's conception of the appropriate notion of a model for a vague language is naturally crucial for one's attitude to the validity of the paradoxical arguments and the claim that I have made about conditionals in vague languages. The question about models could be side-stepped entirely if we adopted a replacement definition of validity. If we say that a schema is valid iff no uniform substitution of nonlogical expressions for its schematic letters yields a sentence which is not true, we have a definition of validity which is as intelligible for vague as for sharp languages. Nevertheless, it is obvious that the old objection to this definition in the case of sharp languages applies too in the case of vague languages: if the background nonlogical vocabulary is rather weak in its expressive power, some invalid schemata may be counted as valid.

This objection is ordinarily taken to motivate the familiar set-theoretic definition of validity. But it does not of course tell against a modal definition of validity which counts a schema as valid iff it is not possible that there be nonlogical vocabulary in some extension of the language which makes the schema not actually true. In fact I have considerable sympathy with the view that not only is this modal definition to be preferred to the set-theoretic one, but that the set-theoretic notions themselves should be explained in modal terms. Nevertheless I shall not discuss validity in vague languages in terms of that modal definition, for what I have to say about such validity should be acceptable to anyone who is able to understand the set-theoretic definition, by whatever route. I shall develop an analogue of the set theoretic conception for vague languages.

A sentence of a sharp language is valid iff it is an instance of a schema that is true under all suitable assignments of sets to its schematic letters. What should stand in the same relation to vague languages as sets thus stand to sharp languages? To play this role I shall introduce the notion of a *sea* of objects. Seas of objects stand to vague predicates as sets stand to classical predicates. There is a binary predicate 'ξ In ζ' true of pairs of objects and seas. This predicate is itself a predicate of degree, and the identity condition of seas is given in terms of it. Where 'a', 'b' ... range over seas, we can say that

S1 $\quad a = b$ iff $\forall x(x \text{ In } a \leftrightarrow x \text{ In } b)$.

Thus seas are 'extensional': if the same objects are in a and b to the same degrees, a is identical with b. Note that this is a sharp identity condition: once the degrees are fixed, either it definitely holds or definitely does not hold. Seas are not 'vague objects' if that phrase is taken to imply that under a given distribution of degrees it can be indeterminate whether the relation of identity holds between a pair of seas.

We can also introduce a sea abstraction operator $\tilde{x}(\ldots . x \ldots)$. An object is in the sea $\tilde{x}(\phi(x))$ to just the degree to which it is ϕ:

S2 $\quad \forall x[x \text{ In } \tilde{x}\ (\phi(x)) \leftrightarrow \phi(x)]$

From S1 and S2 it follows that

S3 $\quad \tilde{x}(\phi(x)) = \tilde{x}(\psi(x))$ iff $\forall x(\phi(x) \leftrightarrow \psi(x))$

This should suffice to convey the fundamental idea of a theory of seas. An axiomatic theory of these entities could be developed, with various existence and comprehension axioms. Provided we operate with degree-functional notions of negation, alternation and conjunction, we can make sense of sea-theoretic operations of complementation, union and intersection respectively.

Employing the notion of a sea, we can then say: a valid schema of a vague language is one which comes out true under all suitable assignments of seas to its schematic letters. A schema is a logical consequence of a set of schemata if all assignments of seas making all elements of the set true also make true the given schema. Note that in these definitions we exploit the qualitative aspect of the notion of the degree to which an object is in a sea.

This model theory takes vagueness very seriously. Not only do the assigned entities, the seas, have their identity conditions given in terms of a vague predicate; we also of necessity use connectives appropriate to a vague language, the degree functional correctives, in giving the model-theoretic account of truth under an assignment. We say for instance that for an atomic sentence *Ft*, and a sequence *s* of objects from model *M*: true $(\ulcorner Ft \urcorner, M, s) \leftrightarrow s(t)$ In $M(F)$. Here $M(F)$ is *M*'s assignment to *F* and $s(t)$ is *s*'s assignment to the term *t*: the important point is that '↔' is our degree-functional biconditional. (Truth in a model would of course be truth in that model relative to all sequences of objects from the domain of that model.) A similar point holds for the other degree-functional connectives. From this point on it is a mechanical matter to construct the model theory, by mimicking the classical forms with the alterations we have indicated. There is though one final *caveat*: if one is considering a language with more than one vague predicate, one must not assign seas of a uniform kind to different predicates unless one is prepared to accept the consequence that it makes sense to compare the degrees to which those different predicates apply to an object.

NOTES

1 See Dummett (1975). I presuppose familiarity with his argument.

2 Wright (1976), p. 245. It should be emphasized that this is not the only reaction consistent with Wright's arguments. Another option is to hold that we need to change our views about the ways in which governing rules can be identified; yet a third is to suggest that it is not impossible to make coherent use of semantically incoherent expressions.

3 This metalinguistic paradox raises a question for Dummett too: would he wish to say that the predicate *C* is incoherent?

4 Platts (1979), chapter IX

[5] *Ibid*, p. 230.
[6] *Ibid.*, p. 230.
[7] Platts does say that "indiscriminability from an instance of red justifies us, other things being equal, in saying that an item is red" (Platts, 1979, p. 231) and says that other things are not equal when I have been persuaded by a Sorites argument (or chain, of corresponding instances) into calling a manifestly orange object 'red'. But the supporter of the incoherence thesis will just say that this means that observational judgements take precedence over conclusions reached by valid means from true premises: that is, it is already revisionary, and is consistent with the incoherence thesis.
[8] Wright (1976), section IV.
[9] The problem is that we cannot say: particulars x and y are red to the same degree iff $\sim\Diamond\exists z(z \text{ matches } x \;\&\sim z \text{ matches } y)$. For we would need to add that in this possible circumstance the degrees of redness of x and y do not differ from their actual values: and thus the condition becomes circular. A similar problem arises with counterfactuals: for a parallel reason we cannot explain the matching of two objects as their being indiscriminable were they to be juxtaposed. We return to some of these issues in a later section.
[10] We can note, too, that this relativized major premise just displayed bears on Wright's arguments (1976) that the utility and point of vague predicates would disappear if we made them precise. His arguments show that precise predicates could not be applied on the basis of casual observation. I agree. But I disagree with his view that his arguments provide reasons for believing the unrelativized major premise of the Sorites paradox. Since two objects a and b can be observationally indistinguishable and yet be red to different degrees, we cannot conclude from Wright's arguments that in such a case 'b is red' must be just as true as is 'a is red'. All we can conclude is that any difference in degree of truth of these two sentences must correspond to a difference in degrees of redness not detectable by causal observation. But there are such differences.
[11] Wright holds that if we introduce degrees of application, the paradox can still be stated for the predicate "it is on balance justified to predicate 'red' of ξ" (1976, p. 239). If this is not a predicate of degree, I would treat it as I have treated C; if it is a predicate of degree, I would treat it as I shall go on to treat 'red'.
[12] Dummett (1973), pp. 446–447.
[13] Dummett (1975), p. 316.
[14] It is not everywhere clear how to distribute temporal indices in Dummett's argument, and some may wonder if there is a resulting fallacy. This is not so – and in any case the point can be made in a spatial example with a disc:

The colours of a, b and c are as described in the text earlier. The solid straight line is the only visible boundary in the disc. (One can also reproduce the argument for a patch slowly changing in colour.)
[15] Dummett (1979), p. 9
[16] Dummett (1976), p. 97.
[17] One could also reach the conclusion of this paragraph by considering the hypothesis that the shades of all objects are vibrating simultaneously between two indiscriminably different shades.

[18] I have taken it not to be satisfactory to avoid the indeterminacies by appealing to physical magnitudes which are the ground in the objects of their secondary qualities. Such an appeal would not be avoiding indeterminacy by appeal to the nature of the experiences themselves.

BIBLIOGRAPHY

Dummett, Michael: 1973, *Frege: Philosophy of Language*, Duckworth, London.

Dummett, Michael: 1975, 'Wang's paradox', *Synthese* **30**, 301–24. Page references are to this printing: the paper is reprinted in his *Truth and Other Enigmas*, Duckworth, London, 1978.

Dummett, Michael: 1976, 'What is a theory of meaning? (II)' G. Evans and J. McDowell, eds. *Truth and Meaning*, Clarendon Press, Oxford.

Dummett, Michael: 1979, 'Common sense and physics', in G. McDonald, ed., *Perception and Identity*, Macmillan, London.

Platts, Mark: 1979, *Ways of Meaning*, Routledge, London.

Wright, Crispin: 1976, 'Language-mastery and the Sorites paradox', in G. Evans and J. McDowell, eds., *Truth and Meaning*, Clarendon Press, Oxford.