

# 1

## *The phenomena of vagueness*

### I. CENTRAL FEATURES OF VAGUE EXPRESSIONS

The parties to the vigorous debates about vagueness largely agree about which predicates are vague: paradigm cases include ‘tall’, ‘red’, ‘bald’, ‘heap’, ‘tadpole’ and ‘child’. Such predicates share three interrelated features that intuitively are closely bound up with their vagueness: they admit borderline cases, they lack (or at least apparently lack) sharp boundaries and they are susceptible to sorites paradoxes. I begin by describing these characteristics.

Borderline cases are cases where it is unclear whether or not the predicate applies. Some people are borderline tall: not clearly tall and not clearly not tall. Certain reddish-orange patches are borderline red. And during a creature’s transition from tadpole to frog, there will be stages at which it is a borderline case of a tadpole. To offer at this stage a more informative characterisation of borderline cases and the unclarity involved would sacrifice neutrality between various competing theories of vagueness. Nonetheless, when Tek is borderline tall, it does seem that the unclarity about whether he is tall is not merely epistemic (i.e. such that there is a fact of the matter, we just do not know it). For a start, no amount of further information about his exact height (and the heights of others) could help us decide whether he *is* tall. More controversially, it seems that there is no fact of the matter here about which we are ignorant: rather, it is *indeterminate* whether Tek is tall. And this indeterminacy is often thought to amount to the sentence ‘Tek is tall’ being neither true nor false, which violates the classical principle of bivalence. The law of excluded middle may also come into question when we consider instances such as ‘either Tek is tall or he is not tall’.

Second, vague predicates apparently lack well-defined extensions. On a scale of heights there appears to be no sharp boundary between

the tall people and the rest, nor is there an exact point at which our growing creature ceases to be a tadpole. More generally, if we imagine possible candidates for satisfying some vague  $F$  to be arranged with spatial closeness reflecting similarity, no sharp line can be drawn round the cases to which  $F$  applies. Instead, vague predicates are naturally described as having fuzzy, or blurred, boundaries. But according to classical logic and semantics all predicates have well-defined extensions: they cannot have fuzzy boundaries. So again this suggests that a departure from the classical conception is needed to accommodate vagueness.

Clearly, having fuzzy boundaries is closely related to having borderline cases. More specifically, it is the *possibility* of borderline cases that counts for vagueness and fuzzy boundaries, for if all actually borderline tall people were destroyed, 'tall' would still lack sharp boundaries. It might be argued that for there to be no sharp boundary between the  $F$ s and the not- $F$ s just *is* for there to be a region of possible borderline cases of  $F$  (sometimes known as the penumbra). On the other hand, if the range of possible borderline cases between the  $F$ s and the not- $F$ s was itself sharply bounded, then  $F$  would have a sharp boundary too, albeit one which was shared with the borderline  $F$ s, not with the things that were definitely not  $F$ . The thought that our vague predicates are not in fact like this – their borderline cases are not sharply bounded – is closely bound up with the key issue of higher-order vagueness, which will be discussed in more detail in §6.

Third, typically vague predicates are susceptible to sorites paradoxes. Intuitively, a hundredth of an inch cannot make a difference to whether or not a man counts as tall – such tiny variations, undetectable using the naked eye and everyday measuring instruments, are just too small to matter. This seems part of what it is for 'tall' to be a *vague* height term lacking sharp boundaries. So we have the principle [S<sub>1</sub>] if  $x$  is tall, and  $y$  is only a hundredth of an inch shorter than  $x$ , then  $y$  is also tall. But imagine a line of men, starting with someone seven feet tall, and each of the rest a hundredth of an inch shorter than the man in front of him. Repeated applications of [S<sub>1</sub>] as we move down the line imply that each man we encounter is tall, however far we continue. And this yields a conclusion which is clearly false, namely that a man less than five feet tall, reached after three thousand steps along the line, is also tall.

Similarly there is the ancient example of the heap (Greek *soros*, from which the paradox derives its name). Plausibly, [S<sub>2</sub>] if  $x$  is a heap of sand, then the result  $y$  of removing one grain will still be a heap – recognising the vagueness of ‘heap’ seems to commit us to this principle. So take a heap and remove grains one by one; repeated applications of [S<sub>2</sub>] imply absurdly that the solitary last grain is a heap. The paradox is supposedly owed to Eubulides, to whom the liar paradox is also attributed. (See Barnes 1982 and Burnyeat 1982 for detailed discussion of the role of the paradox in the ancient world.)

Arguments with a sorites structure are not mere curiosities: they feature, for example, in some familiar ethical ‘slippery slope’ arguments (see e.g. Walton 1992 and Williams 1995). Consider the principle [S<sub>3</sub>] if it is wrong to kill something at time  $t$  after conception, then it would be wrong to kill it at time  $t$  minus one second. And suppose we agree that it is wrong to kill a baby nine months after conception. Repeated applications of [S<sub>3</sub>] would lead to the conclusion that abortion even immediately after conception would be wrong. The need to assess this kind of practical argumentation increases the urgency of examining reasoning with vague predicates.

Wright (1975, p. 333) coined the phrase *tolerant* to describe predicates for which there is ‘a notion of degree of change too small to make any difference’ to their applicability. Take ‘[is] tall’ (for simplicity, in mentioning predicates I shall continue, in general, to omit the copula). This predicate will count as tolerant if, as [S<sub>1</sub>] claims, a change of one hundredth of an inch never affects its applicability. A tolerant predicate must lack sharp boundaries; for if  $F$  has sharp boundaries, then a boundary-crossing change, however small, will always make a difference to whether  $F$  applies.<sup>1</sup> Moreover, a statement of the tolerance of  $F$  can characteristically serve as the inductive premise of a sorites paradox for  $F$  (as in the example of ‘tall’ again).

Russell provides one kind of argument that predicates of a given class are tolerant: if the application of a word (a colour predicate, for example) is paradigmatically based on unaided sense perception, it surely cannot be applicable to only one of an indiscriminable pair (1923, p. 87). So such ‘observational’ predicates will be tolerant with

<sup>1</sup> Note that throughout this book, when there is no potential for confusion I am casual about omitting quotation marks when natural language expressions are not involved, e.g. when talking about the predicate  $F$  or the sentence  $p \ \& \ \neg p$ .

respect to changes too small for us to detect. And Wright develops, in detail, arguments supporting the thesis that many of our predicates are tolerant (1975 and 1976). In particular, consideration of the role of ostension and memory in mastering the use of such predicates appears to undermine the idea that they have sharp boundaries which could not be shown by the teacher or remembered by the learner. Arguments of this kind are widely regarded as persuasive: I shall refer to them as ‘typical arguments for tolerance’. A theory of vagueness must address these arguments and establish what, if anything, they succeed in showing, and in particular whether they show that the inductive premise of the sorites paradox holds.

Considerations like Russell’s and Wright’s help explain why vague predicates are so common (whatever we say about the sorites premise). And they also seem to suggest that we *could* not operate with a language free of vagueness. They make it difficult to see vagueness as a merely optional or eliminable feature of language. This contrasts with the view of vagueness as a defect of natural languages found in Frege (1903, §56) and perhaps in Russell’s uncharitable suggestion (1923, p. 84) that language is vague because our ancestors were lazy. A belief that vagueness is inessential and therefore unimportant may comfort those who ignore the phenomenon. But their complacency is unjustified. Even if we could reduce the vagueness in our language (as science is often described as striving to do by producing sharper definitions, and as legal processes can accomplish via appeal to precedents), our efforts could not in practice eliminate it entirely. (Russell himself stresses the persistent vagueness in scientific terms, p. 86; and it is clear that the legal process could never reach absolute precision either.) Moreover, in natural language vague predicates are ubiquitous, and this alone motivates study of the phenomenon irrespective of whether there could be usable languages entirely free of vagueness. Even if ‘heap’ could be replaced by some term ‘heap★’ with perfectly sharp boundaries and for which no sorites paradox would arise, the paradox facing our actual vague term would remain.<sup>2</sup> And everyday reasoning takes place in vague language, so no account of good ordinary reasoning can ignore vagueness.

<sup>2</sup> See Carnap 1950, chapter 1, Haack 1974, chapter 6 and Quine 1981 on the replacement of vague expressions by precise ones, and see Grim 1982 for some difficulties facing the idea. Certain predicates frequently prompt the response that there *is* in fact a sharp boundary for their strict application, though we use them more loosely – in par-

In the next section I shall discuss the variety of vague expressions – a variety which is not brought out by the general form of arguments for tolerance. First, I clarify the phenomenon by mentioning three things that vagueness in our sense (probably) is not.

(a) The remark ‘Someone said something’ is naturally described as vague (who said what?). Similarly, ‘ $X$  is an integer greater than thirty’ is an unhelpfully vague hint about the value of  $X$ . Vagueness in this sense is underspecificity, a matter of being less than adequately informative for the purposes in hand. This seems to have nothing to do with borderline cases or with the lack of sharp boundaries: ‘is an integer greater than thirty’ has sharp boundaries, has no borderline cases, and is not susceptible to sorites paradoxes. And it is not because of any possibility of borderline people or borderline cases of saying something that ‘someone said something’ counts as vague in the alternative sense. I shall ignore the idea of vagueness as underspecificity: in philosophical contexts, ‘vague’ has come to be reserved for the phenomenon I have described.

(b) Vagueness must not be straightforwardly identified with paradigm context-dependence (i.e. having a different extension in different contexts), even though many terms have both features (e.g. ‘tall’). Fix on a context which can be made as definite as you like (in particular, choose a specific comparison class, e.g. current professional American basketball players): ‘tall’ will remain vague, with borderline cases and fuzzy boundaries, and the sorites paradox will retain its force. This indicates that we are unlikely to understand vagueness or solve the paradox by concentrating on context-dependence.<sup>3</sup>

(c) We can also distinguish vagueness from ambiguity. Certainly, terms can be ambiguous *and* vague: ‘bank’ for example has two quite different main senses (concerning financial institutions or river edges), both of which are vague. But it is natural to suppose that ‘tadpole’ has a univocal sense, though that sense does not determine a sharp, well-defined extension. Certain theories, however, do

ricular, strictly no one is bald unless they have absolutely no hair (see e.g. Sperber and Wilson 1986). But even if this line is viable in some cases, it is hopeless for the majority of vague predicates. E.g. should someone count as ‘tall’ only if they are as tall as possible? How about ‘quite tall’? Or ‘very hairy’? And where is the strict boundary of ‘chair’?

<sup>3</sup> There have, however, been some attempts at this type of solution to the sorites paradox using, for example, more elaborate notions of the context of a subject’s judgement (see e.g. Raffman 1994).

attempt to close the gap between vagueness and a form of ambiguity (see chapter 7, §1).

## 2. TYPE OF VAGUE EXPRESSIONS

So far, I have focused on a single dimension of variation associated with each vague predicate, such as height for ‘tall’ and number of grains for ‘heap’. But many vague predicates are *multi-dimensional*: several different dimensions of variation are involved in determining their applicability. The applicability of ‘big’, used to describe people, depends on both height and volume; and even whether something counts as a ‘heap’ depends not only on the number of grains but also on their arrangement. And with ‘nice’, for example, there is not even a clear-cut set of dimensions determining the applicability of the predicate: it is a vague matter which factors are relevant and the dimensions blend into one another.

The three central features of vague predicates are shared by multi-dimensional ones. There are, for example, borderline nice people: indeed, some are borderline *because* of the multi-dimensionality of ‘nice’, by scoring well in some relevant respects but not in others. Next consider whether multi-dimensional predicates may lack sharp boundaries. In the one-dimensional case,  $F$  has a sharp boundary (or sharp boundaries) if possible candidates for it can be ordered with a point (or points) marking the boundary of  $F$ ’s extension, so that everything that falls on one side of the point (or between the points) is  $F$  and nothing else is  $F$ . For a multi-dimensional predicate, there may be no uniquely appropriate ordering of possible candidates on which to place putative boundary-marking points. (For instance, there is no definite ordering of people where each is bigger than the previous one; in particular, if ordered by height, volume is ignored, and vice versa.) Rather, for a sharply bounded two-dimensional predicate the candidates would be more perspicuously set out in a two-dimensional space in which a boundary could be drawn, where the two-dimensional region enclosed by the boundary contains all and only instances of the predicate. With a *vague* two-dimensional predicate no such sharp boundary can be drawn. Similarly, for a sharply bounded predicate with a clear-cut set of  $n$  dimensions, the boundary would enclose an  $n$ -dimensional region containing all of its instances; and vague predicates will lack such a sharp

boundary.<sup>4</sup> When there is no clear-cut set of dimensions – for ‘nice’, for example – this model of boundary-drawing is not so easily applied: it is then not possible to construct a suitable arrangement of candidates on which to try to draw a boundary of the required sort. But this, I claim, is distinctive of the vagueness of such predicates: they have no sharp boundary, but nor do they have a fuzzy boundary in the sense of a rough boundary-area of a representative space. ‘Nice’ is so vague that it cannot even be associated with a neat array of candidate dimensions, let alone pick out a precise area of such an array.

Finally, multi-dimensional vague predicates are susceptible to sorites paradoxes. We can construct a sorites series for ‘heap’ by focusing on the number of grains and minimising the difference in the arrangement of grains between consecutive members. And for ‘nice’ we could take generosity and consider a series of people differing gradually in this respect, starting with a very mean person and ending with a very generous one, where, for example, other features relevant to being nice are kept as constant as possible through the series.

Next, I shall argue that comparatives as well as monadic predicates can be vague. This has been insufficiently recognised and is sometimes denied. Cooper 1995, for example, seeks to give an account of vagueness by explaining how vague monadic predicates depend on comparatives, taking as a starting point the claim that ‘classifiers in their grammatically positive form [e.g. “large”] are vague, while comparatives are not’ (p. 246). With a precise comparative, ‘*F*-er than’, for any pair of things *x* and *y*, either *x* is *F*-er than *y*, *y* is *F*-er than *x*, or they are equally *F*. This will be the case if there is a determinate ordering of candidates for *F*-ness (allowing ties). For example, there is a one-dimensional ordering of the natural numbers relating to the comparative ‘is a smaller number than’, and there are no borderline cases of this comparative, which is paradigmatically precise. Since ‘is a small number’ is a vague predicate, this shows how vague positive forms can have precise comparatives. It may seem that

<sup>4</sup> Could there be a single, determinate way of balancing the various dimensions of a multi-dimensional predicate that *does* yield a unique ordering? Perhaps, but this will usually not be the case, and when it is, it may then be appropriate to treat the predicate as one-dimensional, even if the ‘dimension’ is not a natural one. Further discussion of this point would need a clearer definition of ‘dimension’, but this is not important for our purposes.

'older than' also gives rise to an ordering according to the single dimension of age, and hence that 'older than' must be precise. But, in fact, there could be borderline instances of the comparative due to indeterminacy over exactly what should count as the instant of someone's birth and so whether it is before or after the birth of someone else. And such instances illustrate that there is *not*, in fact, an unproblematic ordering of *people* for 'older than', even though there is a total ordering of *ages*, on which some people cannot be exactly placed. Similarly, though there is a single dimension of height, people cannot always be exactly placed on it and assigned an exact height. For what exactly should count as the top of one's head? Consequently there may also be borderline instances of 'taller than'.

Comparatives associated with multi-dimensional predicates – for example 'nicer than' and 'more intelligent than' – are typically vague. They have borderline cases: pairs of people about whom there is no fact of the matter about who is nicer/more intelligent, or whether they are equally nice/intelligent. This is particularly common when comparing people who are nice/intelligent in different ways. There are, however, still clear cases of the comparative in addition to borderline cases – it is not that people are *never* comparable in respect to niceness – thus the vague 'nicer than', like 'nice' itself, has clear positive, clear negative and borderline cases.

Can comparatives also lack sharp boundaries? Talk of boundaries, whether sharp or fuzzy, is much less natural for comparatives than for monadic predicates. But we might envisage precise comparatives for which we could systematically set out ordered pairs of things,  $\langle x, y \rangle$  and draw a sharp boundary around those for which it is true that  $x$  is  $F$ -er than  $y$ . For example, if  $F$  has a single dimension then we could set out pairs in a two-dimensional array, where the  $x$  co-ordinate of a pair is determined by the location along the dimension of the first of the pair, and the  $y$  co-ordinate by that of the second. The boundary line could then be drawn along the diagonal at  $x = y$ , where pairs falling beneath the diagonal are definitely true instances of the comparative ' $x$  is  $F$ -er than  $y$ ', and those on or above are definitely false. But for many comparatives, including 'nicer than', there could not be such an arrangement and this gives a sense in which those comparatives lack sharp boundaries.

Another possible sense in which comparatives may lack sharp boundaries is the following. Take the comparative 'redder than' and



choose a purplish-red patch of colour,  $a$ . Then consider a series of orangeish-red patches,  $x_i$ , where  $x_{i+1}$  is redder than  $x_i$ . It could be definitely true that  $a$  is redder than  $x_0$  (which is nearly orange), definitely not true that  $a$  is redder than  $x_{100}$ , where not only are there borderline cases of ' $a$  is redder than  $x_i$ ' between them, but there is no point along the series of  $x_i$  at which it suddenly stops being the case that  $a$  is redder than  $x_i$ . So, certain comparatives have borderline cases and exhibit several features akin to the lack of sharp boundaries: they should certainly be classified as vague.

Having discussed vague monadic predicates and vague comparatives, I shall briefly mention some other kinds of vague expressions. First, there can be other vague dyadic relational expressions. For example, 'is a friend of' has pairs that are borderline cases. Adverbs like 'quickly', quantifiers like 'many' and modifiers like 'very' are also vague. And, just as comparatives can be vague, particularly when related to a multi-dimensional positive, so can superlatives. 'Nicest' and 'most intelligent' have vague conditions of application: among a group of people it may be a vague matter, or indeterminate, who is the nicest or the most intelligent. And vague superlatives provide one way in which to construct vague singular terms such as 'the nicest man' or 'the grandest mountain in Scotland', where there is no fact of the matter as to which man or mountain the terms pick out. Terms with plural reference like 'the high mountains of Scotland' can equally be vague.

A theory of vagueness should have the resources to accommodate all the different types of vague expression. And, for example, we should reject an account of vagueness that was obliged to deny the above illustrated features of certain comparatives in order to construct its own account of vague monadic predicates. (See chapter 5, §2 about this constraint in connection with degree theories.) The typical focus on monadic predicates need not be mistaken, however. Perhaps, as Fine suggests, all vagueness is reducible to predicate vagueness (1975, p. 267), though such a claim needs supporting arguments. Alternatively, vagueness might manifest itself in different ways in different kinds of expression, and this could require taking those different expression-types in turn and having different criteria of vagueness for comparatives and monadic predicates. Another possibility is to treat complete sentences as the primary bearers of vagueness, perhaps in their possession of a non-classical truth-value.

This approach would avoid certain tricky questions about whether the vagueness of a particular sentence is 'due to' a given expression. For example, in a case where it is indeterminate exactly what moment *a* was born and whether it was before the birth of *b*, we would avoid the question whether this shows 'older than' to be vague, or whether the indeterminacy should be put down to vagueness in *a* itself. Provided one can still make sense of a typical attribution of vagueness to some element of a sentence in the uncontroversial cases, I suggest that this strategy is an appealing one.

### 3. VAGUENESS IN THE WORLD?

Is it only linguistic items – words or phrases – that can be vague? Surely not: thoughts and beliefs are among the mental items which share the central characteristics of vagueness; other controversial cases include perceptions. What about the world itself: could the world be vague as well as our descriptions of it? Can there be vague objects? Or vague properties (the ontic correlates of predicates)? Consider Ben Nevis: any sharp spatio-temporal boundaries drawn around the mountain would be arbitrarily placed, and would not reflect a natural boundary. So it may seem that Ben Nevis has fuzzy boundaries, and so, given the common view that a vague object is an object with fuzzy, spatio-temporal boundaries, that it is a vague object. (See e.g. Parsons 1987, Tye 1990 and Zemach 1991 for arguments that there are vague objects.) But there are, of course, other contending descriptions of the situation here. For example, perhaps the only objects we should admit into our ontology are precise/sharp although we fail to pick out a single one of them with our (vague) name 'Ben Nevis'. It would then be at the level of our representations of the world that vagueness came in. (See chapter 7, §1 on an indeterminate reference view.)<sup>5</sup>

My concern is with linguistic vagueness and I shall generally ignore ontic vagueness. This would be a mistake if a theory of linguistic vagueness had to rely on ontic vagueness. But that would be surprising since it seems at least possible to have vague language in a non-vague world. In particular, even if all objects, properties and

<sup>5</sup> The most discussed strand of the ontic vagueness debate focuses on Evans's formal argument which aims to establish a negative answer to his question 'Can there be vague objects?' (1978; see Keefe and Smith 1997b, §5 for an overview of the debate).

facts were precise, we would still have reason, for everyday purposes, to use a vague expression such as ‘tall’, which would still have borderline cases (even if those cases could also be described in non-vague terms involving precise heights etc.). Similarly, in a precise world we would still use vague singular terms, perhaps to pick out various large collections of precise fundamental particulars (e.g. as clouds or mountains) where the boundaries of those collections are left fuzzy. So it seems that language could still be vague if the concrete world were precise.<sup>6</sup>

The theories of vagueness of this book are theories of linguistic vagueness and in the next section I briefly introduce them.

#### 4. THEORIES OF VAGUENESS

The candidate theories of vagueness can be systematically surveyed by considering how they address two central tasks. The first is to identify the logic and semantics for a vague language – a task bound up with providing an account of borderline cases and of fuzzy boundaries. The second task is that of addressing the sorites paradox.

##### (i) *The logic and semantics of vagueness*

The simplest approach is to retain classical logic and semantics. Borderline case predications *are* either true or false after all, though we do not and cannot know which. Similarly, despite appearances, vague predicates have well-defined extensions: there is a sharp boundary between the tall people and the rest, and between the red shades of the spectrum and the other colours. As chapter 3 will describe, the *epistemic view* takes this line and accounts for vagueness in terms of our ignorance – for example, ignorance of where the sharp boundaries to our vague predicates lie. And a *pragmatic account* of vagueness also seeks to avoid challenging classical logic and semantics, but this time by accounting for vagueness in terms of pragmatic relations between speakers and their language: see chapter 6.

<sup>6</sup> These are only *prima facie* reasons for not approaching linguistic vagueness via ontic vagueness: a tighter case would require clarification of what vagueness in the world would be. They also do not seem to bear on the question whether there can be vague sets, which might also be counted as a form of ontic vagueness. Tye, for example, believes that there are vague sets and maintains that they are crucial to his own theory of the linguistic phenomena (see Tye 1990).

If we do not retain classical logic and semantics, we can say instead that when  $a$  is a borderline case of  $F$ , the truth-value of ‘ $a$  is  $F$ ’ is, as Machina puts it, ‘in some way peculiar, or indeterminate or lacking entirely’ (1976, p. 48). This generates a number of non-classical options.

Note that a borderline case of the predicate  $F$  is equally a borderline case of not- $F$ : it is unclear *whether or not* the candidate is  $F$ . This symmetry prevents us from simply counting a borderline  $F$  as not- $F$ . But there are several ways of respecting this symmetry. Some take the line that a predication in a borderline case is both true *and* false: there is a truth-value glut. This can be formalised within the context of a paraconsistent logic – a logic that admits true contradictions (see Hyde 1997 and chapter 7, §7 for discussion of that view).

A more popular position is to admit truth-value *gaps*: borderline predications are neither true nor false. One elegant development is *supervaluationism*. The basic idea is that a proposition involving the vague predicate ‘tall’, for example, is true (false) if it comes out true (false) on all the ways in which we can make ‘tall’ precise (ways, that is, which preserve the truth-values of uncontentionally true or false cases of ‘ $a$  is tall’). A borderline case, ‘Tek is tall’, will be neither true nor false, for it is true on some ways of making ‘tall’ precise and false on others. But a classical tautology like ‘either Tek is tall or he is not tall’ *will* still come out true because wherever a sharp boundary for ‘tall’ is drawn, that compound sentence will come out true. In this way, the supervaluationist adopts a non-classical *semantics* while aiming to minimise divergence from classical *logic*. A theory of this type will be defended in chapters 7 and 8.

Rather than holding that predications in borderline cases lack a truth-value, another option is to hold that they have a third value – ‘neutral’, ‘indeterminate’ or ‘indefinite’ – leading to a three-valued logic (see chapter 4). Alternatively, *degree theories* countenance degrees of truth, introducing a whole spectrum of truth-values from 0 to 1, with complete falsity as degree 0 and complete truth as degree 1. Borderline cases each take some value between 0 and 1, with ‘ $x$  is red’ gradually increasing in truth-value as we move along the colour spectrum from orange to red. This calls for an infinite-valued logic or a so-called ‘fuzzy logic’, and there have been a variety of different versions (see chapter 4).

So far the sketched positions at least agree that there is *some* positive

account to be given of the logic and semantics of vagueness. Other writers have taken a more pessimistic line. In particular, Russell claims that logic assumes precision, and since natural language is not precise it cannot be in the province of logic at all (1923, pp. 88–9). If such a ‘no logic’ thesis requires wholesale rejection of reasoning with vague predicates – and hence of most reasoning in natural language – it is absurdly extreme. And arguments involving vague predicates are clearly not all on a par. For example, ‘anyone with less than 500 hairs on his head is bald; Fred has less than 500 hairs on his head; therefore Fred is bald’ is an unproblematically good argument (from Cargile 1969, pp. 196–7). And, similarly, there are other ways of arguing with vague predicates that should certainly be rejected. Some account is needed of inferences that are acceptable and others that fail, and to search for systematic principles capturing this is to seek elements of a logic of vague language. So, I take the pessimism of the no-logic approach to be a very last resort, and in this book I concentrate on more positive approaches.

Focusing on the question how borderline case predication should be classified, we seem to have exhausted the possibilities. They may be true or false, or have no truth value at all (in particular, being neither true nor false), or be both true and false, or have a non-classical value from some range of values. When it comes to surveying solutions to the sorites paradox, however, there may additionally be alternatives that do not provide a theory of vagueness and perhaps do not answer the question how borderline cases are to be classified. I concentrate on those which do fit into a theory of vagueness.

(ii) *The sorites paradox*

A paradigm sorites set-up for the predicate  $F$  is a sequence of objects  $x_i$ , such that the two premises

- (1)  $Fx_1$
- (2) For all  $i$ , if  $Fx_i$  then  $Fx_{i+1}$

both appear true, but, for some suitably large  $n$ , the putative conclusion

- (3)  $Fx_n$

seems false. For example, in the case of ‘tall’, the  $x_i$  might be the series of men described earlier, each a hundredth of an inch shorter

than the previous one and where  $x_1$  is seven feet tall. (1) ' $x_1$  is tall' is then true; and so, it seems, is the inductive premise, (2) 'for all  $i$ , if  $x_i$  is tall, so is  $x_{i+1}$ '. But it is surely false that (3)  $x_{3000}$  – who is only 4 feet 6 inches – is tall.

A second form of sorites paradox can be constructed when, instead of the quantified inductive premise (2), we start with a collection of particular conditional premises,  $(2C_i)$ , each of the form 'if  $Fx_i$  then  $Fx_{i+1}$ '. For example,

$(2C_1)$  if  $x_1$  is tall, so is  $x_2$

$(2C_2)$  if  $x_2$  is tall, so is  $x_3$

and so on. And the use of conditionals is not essential: we can take a sequence of premises of the form  $\neg(Fx_i \ \& \ \neg Fx_{i+1})$  – a formulation that goes back at least to Diogenes Laertius (see Long and Sedley 1987, p. 222). Alternatively, (2) could be replaced by a quantification over the negated conjunctions of that form.

As well as needing to solve the paradox, we must assess that general form of argument because it is used both in philosophical arguments outside the discussion of vagueness (e.g. with the story of the ship of Theseus) and in various more everyday debates (the slippery slope arguments mentioned in §1).<sup>7</sup>

Responses to a sorites paradox can be divided into four types. We can:

(a) deny the validity of the argument, refusing to grant that the conclusion follows from the given premises; or

(b) question the strict truth of the general inductive premise (2) or of at least one of the conditionals  $(2C_i)$ ; or

(c) accept the validity of the argument and the truth of its inductive premise (or of all the conditional premises) but contest the supposed truth of premise (1) or the supposed falsity of the conclusion (3); or

(d) grant that there are compelling reasons both to take the

<sup>7</sup> As a further example of the former, consider Kirk 1986 (pp. 217ff). Regarding Quine's thesis about the indeterminacy of translation, Kirk uses an argument with the form of the quantificational version of the paradox to argue that there can be no indeterminacy of translation because, first, there would be no indeterminacy in translating between the languages of infants each of whom is at an early stage of language-acquisition and, second, if there is no indeterminacy at one step of acquisition then there is none at the next. He presents his argument as using mathematical induction but does not ask whether its employment of vague predicates casts doubt on that mode of argument.

argument form as valid, and to accept the premises and deny the conclusion, concluding that this demonstrates the incoherence of the predicate in question.

I shall briefly survey these in turn, ignoring here the question whether we should expect a uniform solution to all sorites paradoxes whatever their form and whatever predicate is involved. (Wright 1987 argues that different responses could be required depending on the reasons that support the inductive premise.) Any response must explain away apparent difficulties with accepting the selected solution; for example, if the main premise is denied, it must be explained why that premise is so plausible. More generally, a theory should account for the persuasiveness of the paradox as a paradox and should explain how this is compatible with the fact that we are never, or very rarely, actually led into contradiction.

(a) Denying the validity of the sorites argument seems to require giving up absolutely fundamental rules of inference. This can be seen most clearly when the argument takes the second form involving a series of conditionals, the (2C<sub>i</sub>). The only rule of inference needed for this argument is modus ponens. Dummett argues that this rule cannot be given up, as it is constitutive of the meaning of 'if' that modus ponens is valid (1975, p. 306). To derive the conclusion in the first form of sorites, we only need universal instantiation in addition to modus ponens; but, as Dummett again argues, universal instantiation seems too central to the meaning of 'all' to be reasonably challenged (1975, p. 306). I agree on both points and shall not pursue the matter further here.

There is, however, a different way of rejecting the validity of the many-conditionals form of the sorites. It might be suggested that even though each step is acceptable on its own, chaining too many steps does not guarantee the preservation of truth if what counts as preserving truth is itself a vague matter. (And then the first form of sorites could perhaps be rejected on the grounds that it is in effect short hand for a multi-conditional argument.) As Dummett again notes, this is to deny the transitivity of validity, which would be another drastic move, given that chaining inferences is normally taken to be essential to the very enterprise of proof.<sup>8</sup>

<sup>8</sup> But see Parikh 1983. In my chapter 4, §7 the possibility is briefly entertained.

Rather than questioning particular inference rules or the ways they can be combined, Russell's global rejection of logic for vague natural language leads him to dismiss 'the old puzzle about the man who went bald', simply on the grounds that 'bald' is vague (1923, p. 85). The sorites arguments, on his view, cannot be valid because, containing vague expressions, they are just not the kind of thing that can be valid or invalid.

(b) If we take a formulation of the paradox that uses negated conjunctions (or assume that 'if' is captured by the material conditional), then within a classical framework denying the quantified inductive premise or one of its instances commits us to there being an  $i$  such that ' $Fx_i$  and not- $Fx_{i+1}$ ' is true. This implies the existence of sharp boundaries and the epistemic theorist, who takes this line, will explain why vague predicates appear not to draw sharp boundaries by reference to our ignorance (see chapter 3).

In a non-classical framework there is a wide variety of ways of developing option (b), and it is not clear or uncontroversial which of these entail a commitment to sharp boundaries. For example, the supervaluationist holds that the generalised premise (2) 'for all  $i$ , if  $Fx_i$  then  $Fx_{i+1}$ ' is false: for each  $F^*$  which constitutes a way of making  $F$  precise, there will be *some*  $x_i$  or other which is the last  $F^*$  and is followed by an  $x_{i+1}$  which is not- $F^*$ . But since there is no particular  $i$  for which ' $Fx_i$  and not- $Fx_{i+1}$ ' is true – i.e. true however  $F$  is made precise – supervaluationists claim that their denial of (2) does not mean accepting that  $F$  is sharply bounded (see chapter 7). And other non-classical frameworks may allow that (2) is not true, while not accepting that it is *false*. Tye 1994, for example, maintains that the inductive premise and its negation both take his intermediate truth-value, 'indefinite'.

Degree theorists offer another non-classical version of option (b): they can deny that the premises are strictly true while maintaining that they are *nearly* true. The essence of their account is to hold that the predications  $Fx_i$  take degrees of truth that encompass a gradually decreasing series from complete truth (degree 1) to complete falsity (degree 0). There is never a substantial drop in degree of truth between consecutive  $Fx_i$ ; so, given a natural interpretation of the conditional, the particular premises 'if  $Fx_i$  then  $Fx_{i+1}$ ' can each come out at least very nearly true, though some are not completely true. If the sorites argument based on many conditionals is to count as strictly valid, then



an account of validity is needed that allows a valid argument to have nearly true premises but a false conclusion. But with some degree-theoretic accounts of validity, the sorites fails to be valid – thus a degree theorist can combine responses (a) and (b) (see chapter 4, §7).

Intuitionistic logic opens up the possibility of another non-classical position that can respond to the sorites by denying the inductive premise (2), while not accepting the classical equivalent of this denial,  $(\exists x_i)(Fx_i \ \& \ \neg Fx_{i+1})$ , which is the unwanted assertion of sharp boundaries. Putnam 1983 suggests this strategy. But critics have shown that with various reasonable additional assumptions, other versions of sorites arguments still lead to paradox. In particular, if, as might be expected, you adopt intuitionistic semantics as well as intuitionistic logic, paradoxes recur (see Read and Wright 1985). And Williamson 1996 shows that combining Putnam's approach to vagueness with his epistemological conception of truth still faces paradox. (See also Chambers 1998, who argues that, given Putnam's own view on what would make for vagueness, paradox again emerges.) The bulk of the criticisms point to the conclusion that there is no sustainable account of vagueness that emerges from rejecting classical logic in favour of intuitionistic logic.

(c) Take the sorites (H+) with the premises 'one grain of sand is not a heap' and 'adding a single grain to a non-heap will not turn it into a heap'. If we accept these premises and the validity of the argument, it follows that we will never get a heap, no matter how many grains are piled up: so there are no heaps. Similarly, sorites paradoxes for 'bald', 'tall' and 'person' could be taken to show that there are no bald people, no tall people and indeed no people at all. Unger bites the bullet and takes this nihilistic line, summarised in the title of one of his papers: 'There are no ordinary things' (Unger 1979; see also Wheeler 1975, 1979 and Heller 1988).

The thesis, put in linguistic terms, is that all vague predicates lack serious application, i.e. they apply either to nothing ('is a heap') or to everything ('is not a heap'). Classical logic can be retained in its entirety, but sharp boundaries are avoided by denying that vague predicates succeed in drawing *any* boundaries, fuzzy or otherwise. There will be no borderline cases: for any vague  $F$ , everything is  $F$  or everything is not- $F$ , and thus nothing is borderline  $F$ .<sup>9</sup>

<sup>9</sup> See Williamson 1994, chapter 6, for a sustained attack on various forms of nihilism. For example, he shows how the nihilist cannot state or argue for his own position on

The response of accepting the conclusion of every sorites paradox cannot be consistently sustained. For in addition to (H+), there is the argument (H-) with the premises 'ten thousand grains make a heap' and 'removing one grain from a heap still leaves a heap', leading to the conclusion that a single grain of sand is a heap, which is incompatible with the conclusion of (H+). Such reversibility is typical; given a sorites series of items, the argument can be run either way through them. Unger's response to (H-) would be to deny the initial premise: there are no heaps – as (H+) supposedly shows us – so it is not true that ten thousand grains make a heap. Systematic grounds would then be needed to enable us to decide which of a pair of sorites paradoxes is sound (e.g. why there are no heaps rather than everything being a heap).

Unger is driven to such an extreme position by the strength of the arguments in support of the inductive premises of sorites paradoxes. If our words determined sharp boundaries, Unger claims, our understanding of them would be a miracle of conceptual comprehension (1979, p. 126). The inductive premise, guaranteeing this lack of sharp boundaries, reflects a semantic rule central to the meaning of the vague *F*. But, we should ask Unger, can the tolerance principle expressed in the inductive premise for 'tall' really be more certain than the truth of the simple predication of 'tall' to a seven-foot man? Is it plausible to suppose that the expression 'tall' is meaningful and consistent but that there could not be anything tall, when learning the term typically involves ostension and hence confrontation with alleged examples? A different miracle of conceptual comprehension would be needed then to explain how we can understand that meaning and, in general, how we can use such empty predicates successfully to communicate anything at all. It may be more plausible to suppose that if there are any rules governing the application of 'tall', then, in addition to tolerance rules, there are ones dictating that 'tall' applies to various paradigmatic cases and does not apply to various paradigmatically short people. Sorites paradoxes could then demonstrate the inconsistency of such a set of rules, and this is option (d).

Responses (c) and (d) are not always clearly distinguished. Writers

his own terms (e.g. the expressions he tries to use must count as incoherent). My discussion of methodological matters in chapter 2 will suggest that a swifter rejection of the position is warranted anyway.

like Unger are primarily concerned with drawing ontological conclusions. It is enough for them to emphasise the tolerance of a predicate like 'tall' which already guarantees, they claim, that the world contains nothing that strictly answers to that description: they are not so concerned to examine what further rules might govern the predicate and perhaps render it incoherent. But other writers, for example Dummett, explore these conceptual questions.

(d) Having argued in detail against alternative responses to the paradox, Dummett 1975 maintains that there is no choice but to accept that a sorites paradox for *F* exemplifies an undeniably valid form of argument from what the semantic rules for *F* dictate to be true premises to what they dictate to be a false conclusion. The paradoxes thus reveal the incoherence of the rules governing vague terms: by simply following those rules, speakers could be led to contradict themselves. This inconsistency means that there can be no coherent logic governing vague language.<sup>10</sup>

Once (d)-theorists have concluded that vague predicates are incoherent, they may agree with Russell that such predicates cannot appear in valid arguments. So option (d) can be developed in such a way that makes it compatible with option (a), though this route to the denial of validity is very different from Russell's. (Being outside the scope of logic need not make for incoherence.)

The acceptance of such pervasive inconsistency is highly undesirable and such pessimism is premature; and it is even by Dummett's own lights a pessimistic response to the paradox, adopted as a last resort rather than as a positive treatment of the paradox that stands as competitor to any other promising alternatives. Communication using vague language is overwhelmingly successful and we are never in practice driven to incoherence (a point stressed by Wright, e.g. 1987, p. 236). And even when shown the sorites paradox, we are rarely inclined to revise our initial judgement of the last member of the series. It looks unlikely that the success and coherence in our practice is owed to our grasp of inconsistent rules. A defence of some version of option (a) or (b) would provide an attractive way of

<sup>10</sup> See also Rolf 1981, 1984. Horgan 1994, 1998 advocates a different type of the inconsistency view. He agrees that sorites paradoxes (and other related arguments) demonstrate logical incoherence, but considers that incoherence to be *tempered* or *insulated*, so that it does not infect the whole language and allows us to use the language successfully despite the incoherence (see chapter 8, §2).

escaping the charge of inconsistency and avoiding the extreme, pessimistic strategies of options (c) and (d).

Rather like the liar paradox ('this sentence is false'), where supposed solutions are often undermined by the more resilient 'strengthened liar paradox' (e.g. 'this sentence is false or X' when the response to the original liar is to call it X), a solution to the original sorites paradox can leave untackled other persistent forms of the sorites, or other arguments of a very similar nature. First, consider the phenomenon of higher-order vagueness noted in §1: not only are there no sharp boundaries between the tall and the not-tall, there are no sharp boundaries between the tall and the borderline tall either (see §6). Like the former lack of sharp boundaries, the latter can also be reflected in a sorites premise, e.g. 'growing one thousandth of an inch cannot turn a borderline tall person into a tall one'. Such higher-order paradoxes must also be addressed.

There are also related metalinguistic paradoxes which threaten any theory of vagueness that introduces extra categories for borderline cases assuming they can thereby classify every predication of a given vague predicate in some way or other. In particular, Sainsbury's 'transition question' (1992) and Horgan's 'forced march sorites paradox' (1994) raise similar issues, both emphasising the need to avoid commitment to a sharp boundary between any two types of semantic classification. Horgan instructs us to take, in turn, successive pairs of a sorites series ( $x_1$  and  $x_2$ ,  $x_2$  and  $x_3$  etc.) and report whether they have the same semantic status. If the answer is 'no' for some particular pair then a sharp boundary is drawn between them, contrary to the vague nature of the predicate, but if the answer is always 'yes', all cases will be absurdly classified the same (e.g. the four-foot man will count as tall). And, as Horgan stresses, if a theory commits us to assigning some semantic category to every predication in turn then, assuming they are not all classified the same way, the theory will be stuck on the first horn of his dilemma and committed to sharp boundaries. This emphasises how theorists need to avoid solutions to the original sorites which are still committed to sharp boundaries between semantic categories.

To finish this section I shall briefly mention that there are approaches to the sorites paradox that, I claim, fail to tackle the primary issues

those paradoxes raise. These discussions of the paradox do not slot conveniently into my classification of possible responses, or at least they are not presented as so doing. Unlike most the solutions I have been outlining in (a) to (d), these treatments are not situated in the context of a theory of vagueness more generally. Some, I suggest, may be better seen as tackling a somewhat different issue. For example, sometimes the approach seems to be more of a psychological study of how we respond to successive members of a sorites series and of how our classificatory mechanisms might work such as to prevent us from applying the predicate right through the series. Stories of these kinds do not settle the normative issues of how we *should* classify using vague predicates, what truth-values the problem ascriptions take and what logic governs the language, the very issues I have identified as central to the project in question.<sup>11</sup>

#### 5. THE 'DEFINITELY' OPERATOR

When we construct an account of vagueness, in addition to considering the truth-values of borderline predications, we may seek to *express* the fact that a given predication is or is not of borderline status. Informally, our statement of the fact has relied on semantic ascent – e.g. talking about truth-values of predications. But we may hope to express it without that device. To do so we can introduce into the object language the sentence operators  $D$  and  $I$  such that  $Dp$  holds when  $p$  is determinately or definitely true, and  $I_p$  (equivalent to  $\neg Dp$  &  $\neg D\neg p$ ) holds when  $p$  is indeterminate or borderline. (The terms 'determinately' and 'definitely' are both used in the literature, but

<sup>11</sup> Though I will not argue it here, I consider the treatment in, for example, Raffman 1994 to be of the described type. Among other non-solutions are discussions which give some remedy through which we can avoid actually being driven to paradox (as if it wasn't already clear how this could be done). For example, Shapiro 1998 distinguishes serial processes from parallel ones and attributes the paradox to the use of a serial process that assigns values to predications on the basis of the assignment to the previous member of the ordered sorites series. Such a procedure is wrong because it yields absurd results, Shapiro argues, but he gives no indication of why it is plausible nonetheless (and reliable in other contexts), or what the consequences are regarding sharp boundaries. Moreover, in treating something like the inductive premise of the sorites as an instruction for applying the predicate given certain other members of its extension, Shapiro appears to ignore the fact that it can be treated as a plausible generalisation about the members of the series. On this typical interpretation the paradox persists in abstraction from contexts of running through the sorites sequence via some chosen procedure.

marking no agreed distinction.<sup>12</sup>) This is comparable to the introduction of the sentence operators  $\Box$  and  $\Diamond$  in modal logics: these operators allow an object-language reflection of the meta-linguistic device employed when we report on whether a sentence is possibly or necessarily true. And just as  $\Box$  and  $\Diamond$  can be straightforwardly iterated to express, for example, that necessarily possibly  $p$ , the  $D$  and  $I$  operators can be iterated, where this iteration could perhaps be employed to express higher-order vagueness. For, just as we want to admit borderline cases of  $F$ , where  $\neg DFx \ \& \ \neg D\neg Fx$ , we may want to allow borderline cases of ‘definitely  $F$ ’, where we will have  $\neg DDFx \ \& \ \neg D\neg DFx$ . So one motivation for introducing the  $D$  operator is for the treatment of higher-order vagueness, the issue to which I turn in §6.

We may hold that no sentence can be true without being determinately true. For how can  $a$  be  $F$  without being determinately  $F$ ?  $Dp$  and  $p$  will then be true in exactly the same situations. But the operator is not thereby redundant: for example,  $\neg Dp$  will be true in a borderline case, when  $\neg p$  is indeterminate. When there is some deviation from classical logic and semantics, the fact that  $p$  and  $Dp$  coincide in the way described does not guarantee that they are equivalent in the embedded contexts generated by negating them. (According to the epistemic view, which allows no deviation from classical logic,  $p$  can be true without  $Dp$  being true, namely when  $p$  is borderline and not known to be true. For the  $D$  operator must, on that account, be an epistemic operator.)

The degree theorist can say that  $Dp$  is true if  $p$  is true to degree 1 and is false if  $p$  is true to any lesser degree. A supervaluationist, on the other hand, will say that  $Dp$  is true just in case  $p$  is true on all ways of making it precise and is false otherwise (so if  $p$  is borderline,  $p$  itself will be neither true nor false, but  $Dp$  will be false). Ways of making the whole language precise each yield a model of the language, and definite truth, as truth on all models, may be expected to share structural and logical features with *necessary* truth construed as truth in all worlds.<sup>13</sup> Alternatively the  $D$  operator could perhaps be taken as

<sup>12</sup> Some authors use  $\Delta$  and  $\nabla$  in place of  $D$  and  $I$ , others use *Def* or *Det* for  $D$ ; and some chose an operator for ‘definite *whether*’ (i.e.  $Dp \vee D\neg p$  in my terms).

<sup>13</sup> See chapter 8, §3. Note that it is not only on the supervaluationist scheme that the comparison with modal logics is appropriate. Williamson, for example, explains its applicability within an epistemic view of vagueness (see especially his 1999).

primitive, in the sense that there is no account of it that is derivative from other resources used in a theory of vagueness.

Wright claims that ‘when dealing with vague expressions, it is essential to have the expressive resources afforded by an operator expressing *definiteness* or *determinacy*’ (1987, p. 262). I take this to imply that we will fail to fulfil the central tasks of a theory of vagueness unless we introduce the *D* operator. It is only when we have that operator that we can state that borderline cases occupy a gap between definite truth and definite falsity without committing ourselves to a gap between truth and falsity (Wright 1995, p. 142). And, Wright also maintains, we need to use the *D* operator to say what it is for a predicate to lack sharp boundaries. Consider a series of objects  $x_i$  forming a suitable sorites series for *F* (e.g. our line of men of decreasing heights for ‘tall’). Wright proposes (1987, p. 262)

(W) *F* is not sharply bounded when there is no *i* for which  $DFx_i$   
&  $D\neg Fx_{i+1}$ .

This can be contrasted with the suggestion that a predicate lacks sharp boundaries when there is no *i* such that  $Fx_i$  &  $\neg Fx_{i+1}$ . This latter condition gives rise to paradox; but lacking sharp boundaries in the sense of (W) does *not* lead straight to paradox. In particular, suppose that there are some indefinitely *F* cases between the definitely *F* cases and the definitely not-*F* cases. Then, as (W) requires, there will be no immediate leap from  $DFx_i$  to  $D\neg Fx_{i+1}$  (see also Campbell 1974 and my chapter 7, §5).

Suppose someone were to take Wright’s claim about the importance of *D* to show that a theory of vagueness should proceed by introducing a primitive *D* operator and focusing on its logic and semantics. They would, I argue, be pursuing the wrong approach. Having a primitive *D* operator will not enable us to fulfil the tasks facing a theory of vagueness. In particular, replacing statements naturally used to express our intuitions about borderline cases and the lack of sharp boundaries with *different* (but similar) statements involving the *D* operator does not provide an excuse to ignore the very questions that are, and should be, at the centre of the debate, namely ones about the original intuitions.

For example, suppose the claim is that we are confusing the standard premise with something else which does not lead to paradox, namely the claim that there is no successive pair in the series

of which the first is definitely *F* and the second is definitely not-*F*. We need still to ask how we should classify the original premise itself. If we say that premise is true, as it seems to be, the paradox remains untouched. But can we be content to call it false and hence accept that there is a last patch of a sorites series that is red and an adjacent pair in the series of which one is *F* and the other is not-*F*? Wright suggests that the inductive premises of some (but not all) sorites paradoxes may be of indeterminate status (1987, p. 267). But how are we to understand this claim? At the least it seems to imply that attaching a ‘definitely’ operator to the front of the premise-statement would result in a statement that was not true. But what should we say when we do *not* attach that operator? If regarding that premise as having indeterminate status is to be taken as ascribing it a non-classical truth-value (or just not ascribing a classical value), then a non-classical logic and semantics of vagueness needs to be provided to fill out the picture. But then surely providing such a system should be the central task, rather than concentrating on the logic and semantics of the *D* operator, which would then be an optional extra. Similarly for the claim that our intuition that a borderline predication is neither true nor false should be accounted for by the fact that it is actually neither *definitely* true nor *definitely* false. For are we then to say that it *is* either true or false, and if so, how are we to avoid the unwanted consequences of bivalence? And if, instead, it is said to be of indeterminate status, again we will need a logic that can accommodate such a non-classical truth-value status.<sup>14</sup>

In summary, how can it help to add a *D* operator to the language – creating new sentences that may be shown to be unproblematically true or false – when the task is to illuminate the semantics of the old statements which do not contain this operator? Using a *D* operator may allow us to say that it is not definitely the case that *a* is red, but how can this illuminate the semantics of the vague ‘*a* is red’ itself?

It might be suggested that even if introducing the *D* operator does not provide the *key* to a theory of vagueness, such a theory must at least accommodate and give a plausible semantics for the operator.

<sup>14</sup> Could there be a coherent theory that retains bivalence but still maintains that some sentences are of indeterminate status? This would imply that there could be sentences that were both true and indeterminate, which goes against the earlier assumption that no sentence can be true without being determinately true. Williamson 1995 argues that such a theory is not possible unless the indeterminacy is taken to be epistemic.



For, in describing the semantics of our language we should acknowledge that expressions such as ‘borderline’ and ‘definitely’ are part of it. Moreover, the *D* operator enables us to make assertions about candidates for *F*-ness (e.g. borderline cases) as assertions about the things themselves, whereas without the operator we are strangely limited to judgements about the *language* if we are to say anything like what we want to say. But though I agree that there is a pre-theoretic notion of ‘definitely’, we should be wary of constructing an account of *D* via one’s theory and assuming that it corresponds exactly to a pre-theoretic notion (even if the theory appropriately captures vague language without that operator). The ordinary use and apprehension of ‘definitely’ may well not straightforwardly conform to the kind of formal theory of the *D* operator that theorists seek. Intuitions about the operator may be inconsistent (just like those leading to sorites paradoxes). And, anyway, the consequences of the theory of *D* will outstrip the consequences we would expect given only our intuitions about ‘definitely’. We should beware unargued theoretical assumptions that, for example, *D* can be used to capture the vagueness of *any* expression, including ‘*D*’ itself.<sup>15</sup>

It is thus reasonable, and perhaps necessary, to give ‘definitely’ a technical sense that depends on and is dictated by the theory of vagueness offered for the *D*-free part of language. And the theory may dictate that there is some departure from uniformity between the treatment of sentences with the operator and that of those without. For example, according to supervaluationism, though the logic of the *D*-free language is classical, the logical behaviour of the *D* operator has to be non-classical (see chapter 7, §4). So, although an account of the *D* operator may provide further details of a theory of vagueness, it forms the second and less central stage of such a theory. My prime concern is with the first stage: discussion of the second stage needs to be built on my account of the logic and semantics of *D*-free language.

I now turn to higher-order vagueness, in relation to which the *D* operator remains highly relevant.

<sup>15</sup> Wright offers principles governing the *D* operator, but these are insufficiently defended and are disputed in Sainsbury 1991, Edgington 1993, and Heck 1993. Evans’s celebrated argument concerning indeterminate identity uses a determinately operator (frequently taken to be ‘determinate whether’) and again employs unjustified assumptions about its logic (1978).

6. HIGHER-ORDER VAGUENESS

Imagine – if we can – a predicate  $G$  that has a sharply bounded set of clear positive cases, a sharply bounded set of clear negative cases, and a sharply bounded set of cases falling in between. Although  $G$  is stipulated to have borderline cases in the sense of instances which are neither clearly  $G$  nor clearly not- $G$ , it still has sharp boundaries – one between the  $G$ s and the borderline cases and another between the borderline cases and the not- $G$ s. Our ordinary vague predicates such as ‘tall’, ‘red’ and ‘chair’ surely do not yield a three-fold sharp classification of this sort, with two sharp boundaries around the borderline cases. The familiar arguments that there is no sharp boundary between the positive and negative extensions of ‘tall’ would equally count against any suggestion that there is a sharp boundary between the positive extension and the borderline cases (consider the typical arguments for tolerance discussed in Wright 1976). For example, one hundredth of an inch should not make the difference as to whether someone counts as borderline tall. And a sharp boundary to the borderline cases of  $F$  would mean that there could be two things that are indiscriminable by those who use that word but yet that differ over whether  $F$  applies. More generally, just as the meaning of a vague predicate does not determine a sharp boundary between the positive and negative extensions, nor does it determine sharp boundaries to the borderline cases or other sharp boundaries. (On the epistemic view the requirement would need to be formulated differently, but parallel issues arise; see chapter 3, §1.) With the  $D$  operator, the lack of sharp boundaries to the borderline cases of  $F$  can be expressed as the lack of abrupt transition between the  $DFx$  cases and the  $\neg DFx$  cases (and between the  $D\neg Fx$  cases and the  $\neg D\neg Fx$  cases), or the lack of a last  $x$  in a sorites series for which  $DFx$  is true (and the lack of a first  $x$  in a sorites series for which  $D\neg Fx$  is true).

It is widely recognised in the literature from Russell onwards (1923, p. 87) that the borderline cases of a vague predicate are not sharply bounded. There is disagreement over whether or not a predicate with sharply bounded borderline cases should count as vague (for example, Sainsbury suggests not, 1991, p. 173, in contrast with Fine, 1975, p. 266). But however that question is settled, our ordinary vague predicates typically have borderline cases that are not sharply bounded, so that phenomenon needs to be examined.

Closely related to the lack of sharp boundaries to the borderline cases is the phenomenon of having possible borderline borderline cases (also known as second-order borderline cases), where borderline borderline cases of  $F$  are values of  $x$  for which ‘ $Fx$  is borderline’ is itself borderline. Suppose we accept that the borderline cases of  $H$  are not sharply bounded. We can infer that  $H$  has possible second-order borderline cases given a widely held assumption:

- (A<sub>1</sub>) The lack of sharp boundaries between the  $F$ s and the  $G$ s shows that there are possible values of  $x$  for which ‘ $x$  is borderline  $F$ ’ and ‘ $x$  is borderline  $G$ ’ both hold.

For the lack of a sharp boundary between the definite  $H$ s and the borderline  $H$ s will then imply that there are possible cases between them which are borderline borderline cases of  $H$  as well as borderline cases of ‘definitely  $H$ ’. (And there will be a second variety of possible borderline borderline cases arising from the lack of a sharp boundary between the borderline cases and the definitely false predications.)

This argument for second-order borderline cases looks as though it should now iterate: if  $H$  is to be genuinely vague there should be no sharp boundaries to the borderline borderline  $H$ s either, and this, in turn, will yield possible borderline borderline borderline  $H$ s; and so on. If there is no order of borderline case which we are willing to acknowledge as having sharp boundaries, the iteration will continue indefinitely, resulting in an unlimited hierarchy of possible borderline cases of different orders; we can call this unlimited higher-order vagueness.

The term ‘higher-order vagueness’ has been used for several phenomena which we may wish to keep apart. In particular, sometimes it amounts to having borderline cases of any order above the first; sometimes the term is used to refer to the lack of sharp boundaries to the borderline cases; and occasionally it is used to mean the same as my ‘unlimited higher-order vagueness’. When it is not important to make these distinctions, I shall use ‘higher-order vagueness’ for this cluster of phenomena, though elsewhere it will be preferable to use descriptions without this potential ambiguity (such as ‘the lack of sharp boundaries to the borderline cases’).<sup>16</sup>

<sup>16</sup> Williamson 1999 defends a different characterisation of the hierarchy of orders of vagueness such that there can be third-order vagueness in his sense without third-order borderline cases in the above sense. The dispute does not matter for our current

Another argument for unlimited higher-order vagueness can be constructed as follows. When  $F$  is vague, typically the predicate ‘is a borderline case of  $F$ ’ is also vague. Given the assumption that if a predicate is vague, then it has possible borderline cases – call it  $(A_2)$  – it follows from the vagueness of  $H$  that there are possible borderline cases of  $H$ , since the borderline cases of ‘is borderline  $H$ ’ are themselves borderline borderline cases of  $H$ . And accepting ‘if  $F$  is vague then “is a borderline  $F$ ” is also vague’ would guarantee that if a predicate is vague at all, it has possible borderline cases of all orders. Moreover, since the first-order borderline cases of one predicate coincide exactly with the second-order borderline cases of another, this suggests another (plausible) requirement, namely that higher-order borderline cases be given the same treatment as first-order ones: there should be consistency and uniformity in the treatment of different orders of vagueness.

Should we accept the commitment to an unlimited hierarchy of orders of borderline case associated with each of our typical everyday predicates?<sup>17</sup> It might be thought an extravagant and unrealistic commitment. Moreover, the hierarchy of borderline cases may still fail to capture the complete lack of sharp boundaries for  $F$  if there is a sharp boundary between those cases which are borderline cases of some order and those that are, as we might say, absolutely definitely  $F$ . (See e.g. Sainsbury 1990, p. 11.) And, relatedly, the hierarchy will be of limited benefit if it is such as to pin every candidate for being  $F$  into exactly one of the orders of borderline case, since again this seems to impose a determinacy where there is none.

$(A_1)$  and  $(A_2)$ , the principles used above in generating the hierarchy, both reflect the common emphasis on borderline cases, which seems reasonable at the first level, but may be less compelling once they are seen to draw us into the hierarchy.  $(A_2)$  amounts to the standard criterion of vagueness in terms of borderline cases (a criterion

purposes. Also relevant to a detailed discussion of the various related phenomena would be Williamson’s argument that if  $F$  is second-order vague, it must be vague at all orders.

<sup>17</sup> Burgess argues against unlimited higher-order vagueness, maintaining that higher-order vagueness terminates at a ‘rather low finite level’ (1990, p. 431), at least for secondary-quality predicates. He does this via a proposed analysis of one of the relevant notions, each of the elements of which, he argues, is only vague to a finite level. The strength of his arguments must rest in their detail – for instance they must avoid the objection that precision is simply assumed at some key stage of the account (see Williamson 1994, p. 296).

that can be found in Peirce 1902). It is particularly amenable to an iterative structure. To test whether a predicate  $F$  is vague, ask whether it has borderline cases. If so, this yields a set of cases (those borderline cases) and with regard to them we can apply the same test and ask whether they too have borderline cases (which, if so, will give second-order borderline cases of  $F$ ). And so on through the progressive sets of higher-order borderline cases. An alternative criterion that takes vague predicates to be those lacking sharp boundaries would not have the same scope for the generation of a hierarchy of levels. (Testing whether the borderline cases themselves have borderline cases is no longer iterating the test for vagueness at the first level with this new criterion.) For the fact that there are no sharp boundaries to the positive extension is not a feature susceptible to iteration, just as precision, interpreted as the existence of a sharp boundary, leaves no scope for a notion of higher-order precision. Sainsbury 1990, claiming that there is no such thing as an unsharp boundary, identifies the defining feature of vagueness as 'boundarylessness'. He argues that recognising the feature of boundarylessness is essential for a genuine understanding of vagueness and an account of its semantics. At the very least, we should say that it is more important to capture the lack of sharp boundaries to the borderline cases than to focus on the hierarchy of borderline cases, and this may mean not taking borderline cases as the centre of the debate.

A key issue here concerns vagueness in the metalanguage – the language in which we frame our theory and report the borderline status of some predications. If the metalanguage contains the object-language, so that sentences of the object-language are also sentences of the metalanguage, then the metalanguage will be vague given that the object-language is vague. The interesting issues concern whether the proper part of the metalanguage which is not also part of the object-language is also vague. This part will contain all truth-value predicates, plus expressions for the consequence relation etc. Or if the metalanguage is the same language as the object-language, then we can still ask about the cited elements of the language (they are still called upon to talk *about* the language). If these elements were all precise, then the (precise) metalinguistic predicate applicable to all and only those sentences of borderline case status would pick out a sharply bounded set of cases. But this would guarantee that, for all  $F$ , the borderline cases of  $F$  themselves had sharp boundaries. So

accommodating the lack of such sharp boundaries requires a vague metalanguage. And the existence of higher-order borderline cases would impose the same requirement. For if  $x$  is a borderline borderline case of  $F$ , the metalinguistic report that  $Fx$  is borderline will itself be of borderline status, so the metalinguistic predicate it uses must have borderline cases. Whether and how to accommodate vague metalanguages is a question for any theory of vagueness, and the need for a vague metalanguage is emphasised in Sainsbury 1990, Tye 1990 and Williamson 1994. And Horgan's forced march paradox described in §4 bears on this issue, for with a non-vague metalanguage we can assume that some semantic status or other will be assigned to each object-language sentence, and, as Horgan argues, this will mean sharp boundaries along the relevant series.

There are difficulties, however, facing the idea of vague metalanguages. In chapter 4, §9 I shall argue that certain theories cannot be consistently defended on the supposition that their metalanguages are vague. But even if the metalanguage for some theory could be vague, the following question arises: can we can succeed in illuminating the vagueness of our language if we need to draw on a metalanguage that itself exhibits vagueness? There is at least a suspicion of circularity or triviality here, which has been alluded to in the literature. And Fine suggests that in constructing and assessing theories of vagueness, we might 'require that the meta-language not be vague, or, at least, not so vague in its proper part as the object-language' (1975, p. 297). This tension between needing and resisting vague metalanguages will be explored in later chapters in the context of specific theories.

If we approach higher-order vagueness by using the  $D$  operator within the object language, can we ignore the vagueness or otherwise of the metalanguage? I think not. With the statement (BB)  $\neg DDp \ \& \ \neg DIp$ , we may be able to express the fact that  $p$  is a second-order borderline case that is not definitely definitely true and not definitely borderline. And (BB) can be unproblematically assigned the value 'true' in a non-vague metalanguage. But when we come to assign truth-values to *all* statements of the object language, we will *still* be required to assess the truth-value of  $p$  itself. Being a second-order borderline case,  $p$  is appropriately called neither 'true' nor 'borderline', so even if the metalanguage has an expression for borderline status, that will not be enough unless that expression is itself vague. So a precise metalanguage cannot capture the truth-value status of a

second-order borderline case, and it is not to the point to note that by using  $D$  and  $I$  we can still express the fact that  $p$  is a second-order borderline case.

In summary, I maintain that any putative theory of vagueness must accommodate the apparent lack of sharp boundaries to the borderline cases, and address the issue of higher-order vagueness. And, relatedly, it must answer the question whether the metalanguage for the theory is vague, while tackling the difficulties facing the chosen answer.

## 2

### *How to theorise about vagueness*

In this chapter I shall examine in more detail the project of constructing a theory of vagueness. The apparent simplicity of its central question, ‘what are the logic and semantics of a vague language?’, masks considerable unclarity in the nature of the project and what would count as success. I shall discuss matters of methodology, the aims and constraints of the project and the standards by which we should judge candidate theories. And in §3 I shall investigate an important distinction between two attitudes to elements of a theory.

#### I. ESTABLISHING A REFLECTIVE EQUILIBRIUM

A theory of vagueness deals with the semantic structure of vague languages and the logical relations that hold between their sentences. It must specify the range of truth-values, or any alternative truth-value status, that a sentence can have. And it should capture the distribution of them among borderline cases, in particular through a sorites series. But, as we saw in chapter 1, §4, it may not be theoretically possible to assign some determinate truth-value status to each member of the series in turn – for that would commit us to sharp boundaries between any two semantic categories. The best description of the distribution of truth-values may have to be of some other form, where, for example, we take a step back from describing assignments case by case and describe the general structure of truth-values (see chapter 8, §1). Similarly, if we require of a theory of vagueness that it specifies truth-conditions for vague sentences, then we must allow such conditions to be stated in vague terms in such a way as not to settle a truth-value status for every sentence.

The theory is also concerned with the logical principles governing our language – with, for example, specifying how the truth-value of a complex sentence is determined by, or otherwise related to, its



component sentences (e.g. by giving recursive truth-clauses for the logical connectives). More specifically, it must identify any special logical features that arise owing to vagueness. Other obvious tasks for a theory include solving the sorites paradox and giving an account of higher-order vagueness.

The methodology by which theories of vagueness should be, and generally are, assessed is best seen in terms of establishing a 'reflective equilibrium'. Theorists should aim to find the best balance between preserving as many as possible of our judgements or opinions of various different kinds (some intuitive and pre-philosophical, others more theoretical) and meeting such requirements on theories as simplicity. And when counter-intuitive consequences do follow, the theorist needs to be able to explain *why* we are inclined to make those judgements that their theory regards as erroneous.

This is a familiar strategy in philosophical theorising. Aristotle aimed to produce theories which preserve 'the truth of all the reputable opinions, . . . or, failing this, of the greater number' (*Nicomachean Ethics* VII. 1 1145b5–6). And Rawls (1971, p. 20) uses the phrase 'reflective equilibrium' to describe 'the process of mutual adjustment of principles and considered judgements'. Closer to our topic is Goodman's pioneering discussion of the justification of deduction and of induction, which he shows to proceed in the manner described (Goodman 1954, pp. 63–4). And Lewis emphasises the central role of such a method within much current philosophy; see his 1983b, p. x, where the strategy is explicitly outlined and endorsed for application in a range of different areas, in particular metaphysics.

It is a holistic method: we assess a theory as a whole, by its overall success, allowing counter-intuitive consequences in one part of the theory for the sake of saved intuitions in another part. I do not assume that the best is good enough: it may be that no extant theory of the relevant phenomenon preserves enough intuitions to be acceptable, though I shall not now enter into the difficult question of how good is good enough.

In §2 I shall elaborate on the body of our intuitions, opinions and judgements that are relevant to the construction of theories of vagueness. That body certainly includes the classification of particular cases with respect to a vague predicate (e.g. we all agree that at 6 feet 8 inches someone is tall). Other elements of our linguistic practice are

relevant, in particular our reasoning: we do not, for example, carry on applying a vague predicate right through its sorites series, and a theory should accommodate that fact. And it is too quick to assume that because we do not assent to  $p$ , the theory should count  $p$  as false. It may be, for example, that we all agree that  $q$  and that  $q$  implies  $p$ ; or we may accept propositions relevantly like  $p$  as true. Or maybe  $p$  can only be denied in a framework which fails to respect many more things we say or which fails to respect the way we reason. And there may be pragmatic explanations of the unassertability of  $p$ . Similarly, the opinions of speakers and their commitment to a given claim may sometimes be better tested indirectly by, for example, checking their response to an argument rather than directly asking them whether they believe it: we are not always the best judges of what our own judgements are.

Not all the relevant judgements are appropriately called ‘pre-philosophical’ (e.g. the opinion that classical logic should not be revised unnecessarily; see §2iii). And various theoretical virtues should also be part of the equation: we may be prepared to deny occasional intuitive judgements for the sake of theoretical benefits such as simplicity. There is no sharp division between ‘common-sense’ opinions and those that are part of our general philosophical views, but no such division is needed since both should be taken into account in producing a theory. (Compare Lewis 1983b, p. x: ‘they are all opinions, and a reasonable goal for a philosopher is to bring them into equilibrium’.) Relatedly, consider the distinction between ‘wide’ and ‘narrow’ reflective equilibrium: in relation to an ethical theory, for example, narrow equilibrium takes into account only our moral opinions (i.e. those directly concerning the subject matter in question), while wide reflective equilibrium also brings into the equation opinions and theories about other matters, e.g. psychological or empirical facts, which might bear on the moral judgements we make. Insofar as such a distinction is applicable here (and the distinction is certainly not clear-cut), it is wide reflective equilibrium that I am recommending: there is no restriction on the types of judgements entering into the equation.

Some of our judgements and opinions will need to be given up and counted as wrong despite initial appearances. The sorites paradox brings out the incompatibility of our intuitive judgements that each of its premises is true, its conclusion false, its inference valid, and yet

its predicate coherent. More generally, our intuitions conflict with one another and are not reliable enough in the problematic cases: there will be no suitable system that accommodates all of them. This also answers the objection that my methodology should (absurdly) count as the ideal ‘theory’ that account generated by simply listing off the relevant opinions: such a list would not be consistent. It also would not constitute a theory of vagueness because it would not answer questions as to the status of borderline cases and the logic etc., for these questions are not settled by the list of our opinions and require some theorising to be employed.

Theorists should be most reluctant to deny very widely held judgements or those held among the experts thought most appropriate to judge the matter in question. (Compare Aristotle’s regard for judgements ‘accepted by everyone or by the majority or by the wise’, *Topics I. 1* 100b21.) And they should have a similar regard for those most deeply held, i.e. that we are least prepared to revise. There will also be intuitions and judgements that are particularly important in the context, namely when giving a theory of vagueness. Certain opinions which could be ignored in another context are crucial because to ignore them would be to ignore a key factor bound up with vagueness. For example, the issue of higher-order vagueness is paramount: any account that fails to accommodate it is likely merely to push the untreated problems of vagueness elsewhere. That is unacceptable even if our intuitions about the higher-order matters seem to be less strongly held because they are of less everyday interest and concern.

Given the described methodology, there is unlikely to be any theory which can be conclusively defended: the strategy invites different equilibria reached by choosing to retain different judgements and justifying the sacrifices by emphasising different gains. And apart from by showing a theory to be inconsistent, there will be no test which will refute a theory by showing its incompatibility with certain apparent truths – any apparent truth on which such a test would need to rest may be denied if this is compensated for by those retained and by other virtues the theory can boast. In assessing some given theory, we should determine the extent and range of our judgements which it is forced to deny to reach its equilibrium. Controversy over the success of different theories can then arise in at least three related ways.

First, there can be disputes about what is in the relevant body of opinions – whether some given opinion is really one that we must attempt to save. It may take a (carefully formulated) questionnaire to discover what the opinions of the folk really are. (And it must not be assumed that the corrupted views of the theorising philosopher reflect the common view.) Then, in some cases, two theorists can agree that there is some relevant judgement that we should try to preserve, but disagree over exactly what its content is. One theorist's presentation of an intuitive judgement can be seen by another as prejudiced by the theory advocated. Take, for example, our intuitions about borderline cases. It might be said that we judge the predications to be semantically indeterminate, and so neither true nor false. But epistemic theorists can object that, as Wright puts it (1995, p. 134), 'the ordinary idea of genuine semantic indeterminacy is not itself a datum, but a proto-*theory* of data', and they will maintain that it is enough for the theory to reflect our ignorance in those cases. Another disagreement over what it takes to preserve an opinion arises when non-classical values are admitted. To capture our intuition that  $p$ , must  $p$  count as completely true or is it enough for it to be non-false or perhaps more than 0.5 degrees true?

Second, even if there were agreement over what judgements should be preserved, there could be disagreement concerning some particular theory over which of those judgements it does and does not preserve. Determining the counter-intuitive consequences of a theory is always a major part of its assessment. And we must be cautious of theories that appear to save the, or some of the, high-profile intuitions (e.g. regarding the law of excluded middle) but that do so in a way that requires the denial of a range of other lower profile, but equally important, intuitions.

Third, if we were to have some theories in front of us, along with a list of their counter-intuitive consequences, there could still be considerable disagreement over which of those theories provides the best fit for our body of opinions and intuitions. For it needs to be settled what costs are incurred by denying particular judgements and what would count as adequate compensation for denying them. Different parties to the debate will inevitably value different opinions differently and the methodology does not solve those disagreements.

In chapters 3 to 6 below I reject a series of theories. In most cases I am considering viable accounts which, for the reasons given above,

are unlikely to be defeated by a single fatal blow. I thus build up, in each case, a powerful string of objections which taken together reveal the theory as an unattractive package in poor agreement with the crucial opinions. This is the most common strategy found in detailed discussions of vagueness. My defence of supervaluationism is in a large part a matter of answering a range of objections that opponents stack up, and making a case for its being in substantially better agreement with pre-theoretical opinion.

The method of reflective equilibrium is primarily a method of evaluating theories. It can, to a minimal extent, also figure in the methodology of constructing theories of vagueness – Goodman, for example, considered the deliberative process of mutual adjustments between rules and accepted inferences. But reflective equilibrium does allow theorists to come up with their theory however they like. (Though the merits of a methodology of construction can only be judged by the success according to the reflective equilibrium criteria of the resulting theories.)

There is, I suggest, no possible alternative methodology. Theorists may not be open about their search for a reflective equilibrium of the kind described, but this merely results in them privileging certain intuitions, opinions or considerations and ignoring others; it does not reveal that they have some better methodology to hand or any way of justifying their selection of the constraints that cannot be violated. The methodology I describe recommends assessing the theory on all of the evidence available. All we have to go on, apart from equally inconclusive theoretical considerations already factored in, is linguistic practice in the form of what we (speakers) say and believe and how we reason. My described methodology cannot ensure that theorists take account of all relevant information, but stressing the absence of a unique, small set of over-riding constraints could encourage better practice.<sup>1</sup>

## 2. THE CONSTRAINTS

In this section I shall discuss some intuitions, judgements and opinions that a theory of vagueness should seek to preserve. I refer to these as ‘constraints’ on the theory and many are found explicitly or implicitly

<sup>1</sup> For arguments that reflective equilibrium is the only rational methodology of philosophical inquiry in general see e.g. DePaul 1998.

in current theorising about vagueness. Recall that those constraints should in general be regarded as defeasible: each could be denied if this were to result in sufficient benefits. My discussion will also illustrate the lack of consensus over appropriate constraints: many debates about them are largely, as Machina describes one of them, just ‘a battle of raw intuitions’ (1976, p. 51).

(i) *Classification of sentences and arguments*

The first type of constraint concerns genuinely pre-philosophical judgements about what claims made in our vague language are true or are false. For example, there would be universal agreement that ‘Todd is a tadpole’ is true when Todd has just been born (even though he will later become a frog) and that ‘one grain of sand makes a heap’ is false. A theory is obliged to respect such judgements. Giving up one or two such judgements would not be unreasonable given that some judgements need to be denied. But no systematic theory of vagueness could be constructed by giving up just a moderate number of such intuitions. The only extant theory that systematically denies them is Unger’s nihilism, which is forced to reject as false a huge range of intuitive judgements made in observational vocabulary, in particular everything of the form ‘ $x$  is  $F$ ’ with atomic, vague  $F$  and any  $x$ .

When the applicability of non-observational predicates is in question, it might be necessary to deny judgements commanding widespread assent – fool’s gold should not count as gold just because a large number of fools think it is gold. But it is another of our intuitions that there is a difference between these cases, even if the distinction between observational and non-observational vocabulary is not easily or precisely drawn. Similarly, though with observational terms there may still be some systematic errors of judgement (e.g. of someone who looks deceptively tall), a theory should and will be able to explain in intuitive terms the discrepancy in such cases (once people know the nature of the case, they would no longer concur with the original casual judgement).

As well as respecting our strong, agreed opinions about the truth or falsity of particular claims, we might demand that a theory similarly reflects those cases where we hesitate in judging either way or deny both judgements or disagree with each other or change our mind

over time. Borderline case predications are typically like that. The theory may call them neither true nor false, or it may characterise them in some other way which helps explain our behaviour.

In addition to simple predications there will be complex sentences about which we have intuitions. Generalisations such as ‘anyone taller than a tall man is also tall’ are surely true, for example. Relatedly, there are Fine’s penumbral connections which reflect logical relations between indefinite sentences. He claims that, regarding a borderline red–pink blob, we are obliged to respect the truth of ‘if the blob is to be red then it is not to be pink’. And similarly ‘if sociology is to be a science then so is psychology’ is to count as a penumbral truth (1975, p. 276). Such cases constrain the relations between the component sentences (e.g. between the values of ‘the blob is red’ and ‘the blob is pink’), and also the account of ‘if’. And the accounts of other connectives will be constrained by other cases. For example ‘sociology is a science and psychology is not’ must be false. Fine’s supervaluational account meets these challenges and respects such penumbral truths, though inevitably, opponents query whether such supposed truths are really truths at all (e.g. Machina 1976, p. 77, Forbes 1983, p. 244). Somewhat more contrived cases of compelling complex sentences include ‘the extension of “tall” does not have sharp boundaries’ and ‘the borderline cases of “tall” are not sharply bounded’ (the latter raising issues of higher-order vagueness).

We should also consider the sentences that a theory declares *logically* true or logically false, comparing them with the typical opinions about such classifications. A logician’s notion of logical truth may not be part of everyday vocabulary, in which case again such opinions will not count as ‘pre-philosophical’. But, first, they will be closely related to common-sense judgements about what could not possibly be false/true. And, second, theories can also be constrained to reflect philosophically informed judgements about logical truth. Relatedly, a theory must make explicit what it takes a logical truth to be. Some of the constraints concerning logical truth will arise from judgements that the standard classification of some given sentence as logically true or logically false should not be challenged by the recognition of vagueness, while others concern features distinctive of vagueness that *are* thought to demand logical revision.

So what sentences should be classified as logically true? One case that is frequently cited is the law of non-contradiction stating that  $\neg(p \ \& \ \neg p)$  is true for all  $p$ . It is said that this should be a principle within the portion of classical logic that goes unchallenged by vagueness (see e.g. Fine 1975, p. 270). Proponents of certain truth-functional many-valued theories acknowledge that this law fails according to their account, noting that its instances are, at least, never completely false. Sometimes they reply – unpersuasively, it seems to me – that this failure is, in fact, a consequence of distinctively vague aspects of language. On the other hand, the failure of the law of excluded middle is often seen as a distinctive characteristic of a language containing vague predicates. It is typically only when we take  $F$  to be precise that we are willing to affirm that, for every individual  $x$ , either  $Fx$  or not- $Fx$ . So some theorists take as a starting point the requirement that a satisfactory theory of vagueness must not classify the classical law of excluded middle as logically true. But again, this is highly controversial, and defenders of the epistemic view and supervaluationism typically offer arguments for maintaining the logical truth of this law (e.g. Fine 1975, pp. 284–6).

Other candidate constraints on a theory of vagueness rest on the classification of particular arguments or types of argument, for there are many intuitive judgements that particular arguments are valid or invalid. A wide range of classically valid arguments seem entirely unthreatened by the recognition of vagueness; for example, why should vagueness threaten the rule of and-elimination? The validity of certain other types of argument is more controversial. For example, Frege noted that contraposition can fail in the presence of vagueness (Frege 1903, p. 65): one case where it looks suspect is in deducing the unacceptable statement ‘if  $a$  is a borderline  $F$  then  $a$  is not  $F$ ’ from the apparently acceptable ‘if  $a$  is  $F$  then  $a$  is not a borderline  $F$ ’. Machina and others have shown how certain classical inferential rules (e.g. *reductio ad absurdum*) fail to be generally valid on the supervaluationist theory (see chapter 7, §4 below for discussion).

We have other intuitions that particular arguments are good or bad in a sense less precise than ‘valid’ and ‘invalid’. In particular, the sorites paradox would be excluded from the list of good arguments (though, as Machina remarks, 1976, p. 75, ‘the common man’ is convinced that “‘slippery slope” arguments are fine if they’re not carried too far’). Additionally, a theory must provide a plausible



definition of validity in terms of the truth-values (and possible truth-values) of premises and conclusions: theories that admit non-classical values or truth-value gaps may need to generalise the classical definition of validity. The classical definition can be expressed in a number of forms which, though classically equivalent, need not coincide in a non-classical framework. An argument is classically valid iff every valuation that makes the premises true also makes the conclusion true, which is equivalent to saying that there is no valuation in which all the premises are true and the conclusion false and also to saying that in any valuation in which none of the premises are false, the conclusion is also not false. With degree theories, for example, each of these conditions will yield different consequence relations, even once the other features of any given system are fixed (see chapter 4, §7). Again, the decision between alternative definitions will require consideration both of the individual arguments that each definition deems good or bad and of general principles and apparent truths. For example, must validity be preservation of *truth* even when the account is non-classical?

(ii) *Some theoretical constraints concerning language-use*

We can construct a very general and abstract restriction on theories of vagueness arising from an idea such as ‘meaning and use are closely related’. In particular, an expression has the meaning it does partly because it is used as it is: use helps to determine, and maybe entirely determines, meaning. No theory of vague language should confer meanings on vague expressions that cannot be reconciled with this strong connection. (See chapter 3 for discussion of such a constraint in relation to the epistemic view.) Similarly, speakers clearly understand vague languages, and no theory of vagueness should imply otherwise. We can require, very loosely, that the theory does not make it seem impossible or extremely unlikely that we would have come to understand vague predicates with the features that theory attributes to them. And we use vague language very successfully: a theory of vagueness must not rule out the possibility of an explanation of how we manage this. Take, for example, Wright’s objection that Dummett’s view of vague predicates as incoherent makes our successful use of vague predicates inexplicable. And Unger complains that if there were sharp bound-

aries to our vague predicates, our understanding of them would be a ‘miracle of conceptual comprehension’, which he takes to be a challenge for those theories committed to such boundaries. But, as I argued in chapter 1, Unger’s own claim that our vague predicates have no positive instances at all seems incompatible with our understanding of them and the fact that we use them successfully to communicate.

A theory of vagueness must be compatible with facts about how vague words come to have the meaning that they have and how those meanings sometimes change. This may involve considerations about the gradual evolution of our language with individual vague words sometimes being added to the vocabulary, and words that are initially very vague sometimes acquiring a less vague sense. And it must allow for the fact that words can be coined, where this process can and often will create a vague expression. For example, Tappenden 1995 envisages a situation in which the US Supreme Court coins the phrase ‘brownrate’ to mean ‘with all deliberate speed’ intending to avoid specifying an exact required speed: such a practice of incomplete stipulation would have the advantage of allowing the extension to become progressively more complete as new circumstances come to light, rather than requiring an exhaustive classification to be fixed at the outset. We can ask of a theory whether it allows for the coinage of vague predicates. Some critics of the epistemic view argue that it fails this test since it must say that if you coin a meaningful word then sharp boundaries are inevitably fixed and no areas of indefiniteness remain (see Tappenden 1995 and Sainsbury 1995b; for a response see Williamson 1997a).

*(iii) The extent of departure from classical logic*

How should the construction of a theory of vagueness be influenced by the idea that classical logic should not be revised, or at least that no such revision should be taken lightly? Williamson writes, ‘Classical logic and semantics are vastly superior to the alternatives in simplicity, power, past success, and integration with theories in other domains’ (1992b, p. 162). But classical logic should not be seen as unrevisable in this realm – the applicability of classical logic is just one more factor whose preservation is desirable, but it has to be balanced against the judgements that it forces us to deny (e.g. the failure of bivalence

in borderline cases and the lack of sharp boundaries to vague predicates). And there can be no consensus over how high we should put the cost of revising classical logic.

First, what exactly is this ‘standard logic’ (as Williamson 1994, p. xi calls it) that we must respect so highly? It would be misleading to suggest that there is a complete system of logic that is appropriate and uncontroversial in the absence of vagueness. We might ask, for example, whether the logic should be free, whether it should be a relevantist logic and what is to be said about liar sentences or sentences involving category mistakes. Such questions may be currently unsettled, but theories of vagueness need not answer them. One way in which such a theory might avoid them entirely would be to provide rules that will modify *any* system of logic and semantics (whether free, relevantist or intuitionistic etc.) into one which can accommodate vagueness. This would permit independence from other logical questions, since, using the rules specified by the theory, a logical system for a vague language could be constructed from any logic that otherwise recommends itself. But this is not the standard form of a theory of vagueness, nor is it clear to what extent it is viable. In general, theorists avoid questions seemingly irrelevant to vagueness by assuming that by ‘classical logic’ we mean first-order predicate calculus with identity and without non-denoting expressions. I shall follow this usage.

Even if we allow that classical logic might be revised, we can still impose the constraint that any departure from classical logic must be well motivated and kept to a minimum. Such a constraint is widely adopted: see for example Fine 1975, p. 286, and Simons (1996, p. 326) states a closely related requirement that we need to ‘rescue what can be rescued of classical logic’. Vagueness should not prompt us to alter logic and semantics any more than is necessary.

One common idea is that situations in which features distinctive of vagueness are absent should display none of the distinctive logical traits introduced to accommodate vagueness, and in those situations there should be no deviation from classical logic. The form of this requirement could differ for different theories. It is perhaps most clear in connection with a theory that introduces an additional non-classical value or values to be assigned to borderline case predications. The constraint then rules that sentences of determinate truth-value status (i.e. taking classical values) must behave classically and stand

in classical relations of entailment etc. to other classically valued sentences. And if a compound sentence has component sentences that are each either completely true or completely false, then the value of the compound must coincide with the value it would receive in the classical framework given the same distribution of values among the component sentences. A constraint of this type is sometimes called a ‘normality’ or ‘fidelity’ constraint, since it requires that in certain special cases, the logic and semantics should be ‘normal’ or ‘faithful’ to the classical norm. For example Machina (1976, p. 55) writes ‘vague propositions sometimes take the classical truth-values . . . and when they do, the usual classical treatment will be just as acceptable for them as it is for precise propositions’.<sup>2</sup>

### 3. MODELS AND ARTEFACTS

In this section, I describe two ways of regarding elements of a theory of vagueness, the realist and the modelling way. The distinction between them cuts across the usual classification of theories of vagueness – the same theory could be defended with either approach. And they are both compatible with the methodology of reflective equilibrium. But my distinction marks a crucial contrast between attitudes to theories that are found in the literature; it is particularly important in relation to degree theories. Making the contrast explicit should also give us a better grasp on the substance of the claim that some logical system is the logic of vagueness and it will have consequences for what count as appropriate ways of assessing proposed theories.

<sup>2</sup> The normality constraint is rarely justified in any detail and should be taken as no more than a defeasible constraint: a theory violating it should not be ruled out. For it constrains truth-values even for *vague* sentences when they take classical values. And could it not be the case, for example, that a vague sub-sentential element of a sentence (a vague predicate, for example) affects the logical behaviour of that whole sentence even if that sentence happens to be classically valued? Enforcing the normality requirement would also prematurely exclude the possibility of a theory that declared true the premises of a sorites, while denying that the conclusion was true (i.e. a theory that gives those sentences the values that they *seem* to take). For if the premises are (classically valued and) true, then the normality condition rules that they must stand in the usual classical relations, hence (given the classical validity of the sorites argument) that the conclusion is also true.

(i) *The distinction*

Suppose a theory of vagueness provides a detailed model of the truth-value status of borderline cases and the inferential features of our vague language. The question arises how much of the model we should take seriously, or take at face value. Regarding a given aspect of the theory, the realist approach takes it at face value, where it is taken to correspond to an aspect of the phenomenon with which the theory is concerned. By contrast, the modeller will take it less seriously and allow that there is nothing corresponding to it. They typically remark of their theory that 'it's only a model' and they use this claim to escape commitments. For example, they might deny that the assignment of truth-values in borderline cases should be viewed in a realist way. You can, of course, be realist about some aspect of the theory and a modeller about others. (I will sometimes talk about the approaches without specifying the elements towards which they are taken, for often either it will be clear which elements are in question, or the context will be such that the modellers fail themselves to specify.)

The realist conception is so called because the descriptions of the language that the theory provides, or commits us to, are intended to describe real features of that language. An epistemic theorist such as Williamson seems to be working with this conception of the project (1992b, 1994). He claims that, together with the facts, the meaning of a sentence (itself determined by its use, or the use of its components) suffices to determine whether it is true or false: all sentences are determined as having one of these two values and the logic governing them is classical.

By contrast, modellers emphasise that giving a theory of vagueness is a matter of modelling our vague language, where there may be elements of the model not corresponding to elements of reality. So, for example, a logical system that works as a model of our language might be committed to certain statements that appear to be descriptive of our language but which should not, in fact, be taken at face value as describing real features of it. According to the modelling approach, accepting a theory with such commitments can be justified if the theory as a whole is successful. For example, a modeller may not take seriously the truth-value ascriptions made in some borderline case predications, while maintaining that treating the language as if all

sentences have some (unique) truth-value from their specified range enables us to draw conclusions about our vague language that we *should* take seriously and at face value. And some of those conclusions drawn (e.g. concerning matters of validity) will illuminate the features of the language distinctive of vagueness, and provide a solution to the sorites paradox. By contrast, for the realist such truth-value ascriptions are to be taken seriously and are to be seen as themselves illuminating vagueness as well as being involved in deriving further informative statements that do so.

Talk of a modelling approach suggests a comparison between the task of producing a theory of vagueness and the enterprise of scientific modelling. Goguen, for example, describes the task as parallel to providing typical mathematical models of empirical phenomena, even taking a theory of vagueness to play comparable predictive roles and to be 'subject to the process of experimental verification and subsequent modification usual in scientific research' (1969, p. 326). But note that it is compatible with the realist conception as well as the modelling approach to regard theories of vagueness as akin to scientific theories construed in terms of models. On the former, we can still see the theory as providing a model; the difference arises over how the model relates to the modelled phenomenon. Realists about *science* are characterised by van Fraassen as seeking a theory for which they can claim 'to have a model which is a faithful replica, in all detail, of our world' (1980, pp. 68–9). For the envisaged realist conception of theories of vagueness, substitute 'language' for 'world' in the quote. What is distinctive of the modelling approach is the fact that replication in every detail is neither expected nor demanded.

I suggest that many degree theorists, in particular, are best seen as adopting the modelling approach towards a number of features of their theories. Goguen, for example, dismisses the belief in 'some Platonic ideal "concepts" or logic embodying their essence' (1969, p. 325), and he emphasises instead that we are constructing models. And degree theorists often treat their intermediate truth-values as values that it is *useful* for us to ascribe to sentences, not as values that the meaning of a sentence (together with the facts) determines to be applicable to that sentence in the circumstances. They typically do not accept that there *is* a unique, exact value applicable to each sentence (in the actual world). Goguen claims (1969, p. 331) that 'any not identically zero function which is continuous decreasing and

asymptotic to zero' would be appropriate for the representation of 'short', implying that there is not a unique function assigning values of shortness that is given by the meaning of 'short', but rather that we can profitably associate any of a range of functions with that predicate. Edgington also writes 'the numbers are to be taken with a pinch of salt . . . But the numbers afford us the luxury of addition, multiplication etc., in exhibiting a model of the structure of . . . reasoning' (Edgington 1992, p. 203; see also her 1997, p. 308, 'there are no exactly correct numbers to assign'). And Machina indicates similar sympathies in claiming that 'fortunately, the assignment of exact values usually doesn't matter much for deciding on logical relations between vague propositions' (Machina 1976, p. 61).

It would, however, also be possible for someone to see a degree theory as uncovering the true structure of our language in accordance with the realist stance and to insist that our sentences *do* each have one, and just one, of an infinite range of numerical degrees of truth. Equally, a proponent of classical logic could try to adopt the modelling approach and argue that it is not that our language has an underlying structure which is classical (as I portrayed Williamson's position), but rather that classical logic is the best, or perhaps the only, satisfactory way to model language. The position which Cargile labels 'nominalistic' may be of this form, and his distinction between this and his 'realistic' position seems to be a special case of my distinction between the modelling and realist approaches. His realistic line takes it to be true that in turning from a tadpole into a frog, a creature, *Amphibius*, ceases to be a tadpole at some precise instant: 'that is how logic requires that change must be, and so it must (logically) be' (Cargile 1969, p. 200). This is contrasted with his nominalistic alternative according to which 'if we are making use of logic, we may be forced to choose some instant arbitrarily to be the instant at which *Amphibius* ceases to be a tadpole, in much the same way that we have to assume, in applying the differential calculus to a physical problem, that matter is infinitely divisible'. The chosen instant will be singled out in the model, but it will be denied that this corresponds to a feature that is privileged in reality. How much sense we can make of that sketchy proposal is unclear: I would argue that to make a genuine theory out of it, the best option is to adopt a supervaluationist account. We would not be led far astray by choosing an arbitrary instant – at least all our inferences not referring to that

instant would be reliable. But to commit ourselves only to truths not resting on the specific choice made, we should quantify over the range of acceptable models, as supervaluationism instructs.

The modelling approach can be pursued in a range of different ways and so a global verdict on them all could be inappropriate. Nonetheless I shall argue that taking that approach all too often amounts to leaving unanswered what seem to be (and what the realist often takes to be) the key questions to be addressed by a theory of vagueness.

*(ii) Artefacts of a model*

We can distinguish between the genuinely representational features of a model which we should take as capturing or revealing aspects of the modelled phenomenon, and those which are mere *artefacts* of the model, not representing anything in the phenomenon itself. The major difference between the realist and the modelling approaches can then be described in terms of the features of a model that are taken to be mere artefacts – the realist approach to a given feature is to take it as representational while the modeller regards it as a mere artefact of the model.

Consider an analogy with the debate in the philosophy of science between van Fraassen's constructive empiricism and his realist opponents (see e.g. van Fraassen 1980). Roughly sketched, the difference between these positions concerns whether descriptions of unobservables (or of relations between observables and unobservables) should be taken at face value as providing true descriptions of features of the world, or whether we should only believe the predictions and descriptions of behaviour at the observable level, though acknowledging that a model involving unobservables can aid understanding and predictions involving observables. The realist about unobservable entities takes the former line and maintains that there are unobservable entities and objective truths about them, while van Fraassen defends the latter line and does not share the belief that there are unobservable entities. He can be regarded as taking statements about unobservable entities to be artefacts of the model which may not correspond to elements of reality, while the realist maintains that they do directly correspond to aspects of reality.

Taking the modelling approach to any type of theory of vagueness



will, I suggest, involve taking truth-value assignments in at least some borderline case predications to be mere artefacts. We have seen that this attitude is common among degree theorists. Consider again the modelling approach towards classical logic. The important contrast with the attitude of the realist who defends classical logic is displayed in the modellers' denial that we need to take sharp boundaries seriously. But if they were to take seriously the truth-value ascriptions in all borderline cases around that boundary, then they would be committed to a realist attitude to the boundaries themselves, whose existence would be guaranteed by those truth-values. So truth-value ascriptions in at least some borderline cases must be regarded as artefacts. A question such as 'what truth-values, if any, are taken by borderline case predications?' – earlier presented as a central question for theories of vagueness – must then be seen by the modeller as sometimes inappropriate. Though values may be assigned, this can only be interpreted instrumentally, which is not to say that they have no truth-value, for that would be to place them, equally mistakenly, in a truth-value gap. But then we need an explanation of why it is not a reasonable question. The modeller talks instead about what values it is useful for us to apply to them in a non-realistically construed theory, but that only raises the questions why such a theory is useful and what features of the modelled phenomena guarantee the utility of a non-realistic model.

Suppose there is information in the model which it is implausible or otherwise undesirable to assume corresponds to elements of the real modelled phenomenon; by adopting the modelling approach, a theorist seems able simply to deny that there is any such correspondence, without this threatening the use of the model. The modeller then avoids a range of objections that the realist must face. For example, first, in discussing the epistemic view in chapter 3, I shall raise a question about what could possibly determine the classical values taken by certain sentences or the sharp boundaries that the epistemic theorist claims the extensions of our vague predicates have. A theory advocating classical logic within the modelling approach need not face these questions: according to such a view the values are simply usefully assigned rather than picking up on elements of reality. Similarly with the demand for the epistemic theorist to explain why we do not know facts in borderline cases and about where sharp boundaries lie: on the realist conception the demand is reasonable,

while on the modelling approach an advocate of classical logic and semantics may hope to deny that there *are* such facts of which we are ignorant. A second example: there is a natural worry about the fact that the degree theorist's model makes assignments of exact numerical values to all sentences, for this seems to impose fine detail and apparent precision that is altogether inappropriate for the modelled vague phenomenon (see chapter 4, §9). By adopting the modelling approach and not taking those assignments seriously, many degree theorists hope to avoid such worries. Similarly, there may be sharp boundaries imposed by a degree theory (e.g. between sentences of degree 1 and those of degree less than 1). The realist would be forced to take them seriously, while on the modelling approach it might be argued that they are mere artefacts of the model.

But unwanted commitments cannot be avoided so easily. It is not acceptable simply to provide a theory and deny that we have to take all of it seriously without explicitly stating what we *are* supposed to take seriously. To justify both defending a model and regarding some of its features as mere artefacts, we need to know what a theory of vagueness is supposedly modelling and what features of the model are genuinely representational. But typically proponents of the modelling approach do not confront this crucial issue: they simply offer a model and choose at will which aspects of it to take seriously, hoping to avoid worries about the other elements by casually remarking that it is 'only a model'. If this is all they do, they are at very best merely gesturing at a theory of vagueness without really providing one. What is needed is an explicit, systematic account of how the model corresponds to or applies to natural language, stating which aspects of the model are representational, and justifying the treatment of others as mere artefacts. It is far from clear how this could be done.

To emphasise this point, I return to the comparison with scientific modelling. We can regard the model itself as an abstract structure (e.g. the mathematical structure defined by the laws of Newtonian mechanics). A scientific theory typically identifies that structure and specifies how it corresponds to the aspect of the world that is being modelled, stating what features of the model reflect the world, or in what respects the model is similar to the world and to what degree. This correspondence is specified by 'application rules' or 'theoretical hypotheses'; for example that 'the positions and velocities of the earth and moon in the earth-moon system are very close to those of a two-

particle Newtonian model with an inverse square central force' (Giere 1988, p. 81). In these terms, my criticism of current theorists within the modelling approach to vagueness is that they offer no application rules, or any indication of an alternative device fulfilling the same crucial role. Realists are not open to a similar objection: in assigning truth-values to sentences and describing logical relations that hold between them, such theorists aim to capture the truth-values that those sentences actually have and the logical relations that they stand in. The ascriptions and descriptions are not a surrogate for capturing some other (unspecified) features of the language; they are to be taken at face value.

To summarise, I object to modellers who attempt to avoid commitments that their models force on them, when, as is usually the case, they state only that their model is not to be taken realistically and not how it *is* to be taken.

There is one way that modellers who are degree theorists might respond in relation to their refusal to take seriously the numerical truth-value assignments. They could maintain that the instantiation of truth-values by sentences forms an ordinal structure, where it is only the ordering of values that counts. Such a structure is legitimately represented numerically, but when we consider that representation we should not take it entirely seriously, since there will be features of the numerical structure (corresponding, e.g., to ratios or intervals between values) with no parallel in the purely ordinal scale. In chapter 5 I shall examine the viability of an account of this type and assess whether it meets worries about the objectionable precision involved in assigning exact values.

But if viable at all, this type of response is only available to a degree theorist. Consider, for example, a theorist who seeks to defend a classical logic theory of vagueness while seeing the resulting models in terms of the modelling approach. With just two truth-values, there is no room for a discrepancy with the structure of the system representing the instantiation of those values in the way that there is in the infinite-valued case, and there will be no non-representational features in the model of the instantiation of those values, in the way that there would be in the numerical representation of an ordinal scale. Similarly for theories committed to exactly three values. Modellers employing these theories would have to deny that truth-value ascriptions in borderline cases should be taken seriously at all,

not just that we should take care with regards to the chosen representation of them. This more extreme type of modelling approach could also be adopted with a degree theory or any other type of theory; the objections I raised above are particularly pressing for such an account. Nonetheless, in the rest of this chapter I shall continue to examine the modelling approach in general, comparing it with the realist approach and considering it in the light of the methodological theses of §§1–2.

(iii) *Models as idealisations*

The modelling approach is frequently combined with the view that theories of vagueness have to be idealisations. Some theorists emphasise that their theory gives a precise model of an imprecise phenomenon, but that this must not be seen as undermining the theory. Rather, they imply, models and systems of logic *must* be precise, so this form of idealisation is unavoidable. Goguen, for example, writes (1969, p. 327), ‘our models are typical purely exact constructions, and we use ordinary exact logic and set theory freely in their development . . . It is hard to see how we can study our subject at all rigorously without such assumptions.’ And Edgington describes her own account as ‘a precise mathematical model of an imprecise phenomenon’, but claims that nonetheless ‘it gives, modulo that imprecision, the structure of the phenomenon. The demand for an exact account of a vague phenomenon is unrealistic. The demand for an account which is precise enough to exhibit its important and puzzling features is not’ (Edgington 1997, p. 305; see also her 1993, p. 200). Thus such modellers maintain that theories of vagueness must be idealisations in virtue of the very nature of the modelled phenomenon, namely vagueness. (Note that the analogy between the modelling approach and instrumentalism in science breaks down here, if not before, but this does not undermine the original comparison since this issue is a matter of the *motivation* for the approach.)

The position of these modellers is thus that we can, at best, produce an idealised model of vague language because all logical systems are precise. But, first, we need to be told why a model must be precise. Is a mathematical model precise simply in virtue of being a *mathematical* structure? Zadeh would say not, since he describes his

own system of fuzzy logic (which uses fuzzy truth-values) as ‘an imprecise logical system’ (1975, p. 407). Alternatively, we might deny that a model is precise in the case where the metalanguage in which it is framed is itself vague. For if the theory must be formulated within a vague language, then the suggestion that the description is of something precise would surely be misplaced and there may be no need for idealisation.

Even if we were to acknowledge the precision of typical models, the modeller’s position on idealisation needs to assume that the logic of vague language cannot be given by a precise system. Machina appears not to share this assumption when he writes, ‘This is not to say that in order to fulfil its mission to handle vague propositions logic itself must become vague. (The study of dead civilizations need not itself be dead.)’ (1976, pp. 47–8). Moreover, some people maintain that being vague is a matter of coming in degrees, and this feature of predicates *can* be captured within a mathematical model. So there may be room for a position which provides a theory of vagueness employing a precise system without yielding to the criticism that it imposes that precision on the phenomenon itself. The claims both that logical systems must be precise and that a precise system is inadequate for an account of vagueness are, at best, too hastily made.

It is correct, though, that if idealisation were necessary in theories of vagueness, we would have to take the modelling approach. From the realist standpoint, an account of an imprecise phenomenon that rendered it precise would be an account that fails, whereas on the modelling approach it is viewed as a reasonable device in modelling a phenomenon. So, if idealisation is unavoidable, we must take the modelling approach if we are to succeed in producing a theory of vagueness at all; and proponents of the modelling approach might be seen as taking as their starting point the inevitable failure to produce a realist account. But adoption of that starting point is unwarranted unless further argument is provided. And it should be a last resort anyway. Compare idealisation in science: if we are interested in movement of objects on a particular real surface, then an idealisation treating it as frictionless may be fine for practical purposes and useful because of its simplicity, but it will never accurately describe actual movement on that surface. Similarly, an idealised model of a language may help for some purposes but it gives the wrong answer to some

questions about the real features of language and in particular to ones bound up with vagueness. Moreover, surely we can still ask about the real phenomenon. There may indeed be a tension between the vagueness of language we are hoping to capture and the use of logical models that have at least an appearance of precision. But to respond by declaring the model to be merely an idealisation is to give up on the task in hand. The way forward is to admit that the metalanguage is itself vague: an accurate account of vagueness may then be given in terms that are themselves vague.

*(iv) Uniqueness*

The two approaches I have been comparing may also differ with respect to the question whether there is a uniquely correct theory of vagueness. Some philosophers imply that there must be a unique system of logic and semantics for vague language when they talk of providing *the* system of logic and semantics (e.g. Fine 1975, p. 297). But might not several systems be equally appropriate? A claim of uniqueness needs to be carefully stated to allow for different non-conflicting systems, such as a propositional logic and a compatible predicate logic, or a system of predicate logic with identity and one without. We can say, at best, that there is a uniquely correct system for a chosen level of generality and stock of logical constants, e.g. a unique propositional logic for vague language.

The realist view that theories uncover aspects of the language strongly suggests, at the very least, restrictions on the range of different systems that could count as giving the logic and semantics of a vague language, and it may ensure there must be a uniquely correct system. In particular, no two theories that disagree over the range of truth-values available for sentences could both be acceptable. And on the assumption that definitions of (at least some of) the connectives are intended to capture natural language connectives, there could not be two acceptable theories that agree on admitting infinitely many values but disagree over the definitions of those connectives and hence over what relations hold between complex sentences and their components.

If we adopt the modelling approach, there are no such immediate reasons to expect a uniquely acceptable logic and semantics for vague language (even for a chosen level of generality). For why think it

would not be possible to construct different models for vague languages that are equally successful? (As Goguen 1969, p. 326, writes: ‘we do not assume there is some unique best theory’.) There is nothing in the conception of the project itself that suggests that uniqueness should be expected, or even desired, even when it comes to deciding the range of truth-values assigned. If the target for a theory is, at least in part, its utility in illuminating the phenomenon without directly corresponding to it, uniqueness should not be assumed – perhaps different ways of modelling vagueness are useful for different purposes and none is ideal for all. The extent of flexibility depends, however, on what aspects of their model the modeller does regard as corresponding to reality. For example, could we be so permissive as to admit that there is no best choice between degree theories and supervaluationism, despite their radically different consequence relations? No theorists seem prepared to accept that.

Moreover, the same considerations implying that uniqueness need not be expected among theories when they are seen as mere modelling devices also suggest that regarding some given theory in that way could be compatible with there being a different theory meeting realist standards. For the correctness of the latter theory does not prevent some other theory from suitably illuminating the phenomenon of vagueness in a different way. This observation bears on my earlier criticisms of certain versions of the modelling approach, namely that they leave unanswered certain key questions (e.g. about real truth-values as opposed to what it is simply useful to treat as truth-values). It is the search for a theory meeting realist standards that really counts, and if we are to settle for a modeller’s offering we need reasons to believe that no realist theory is possible and an explanation of why the questions it confronts are unanswerable.

In §1 I suggested that the holistic method described there encourages the thought that we could reach different equilibria that balance our intuitions in different ways but that seem equally plausible and where each theory counts as best according to slightly different (but equally valid) standards of judging the ‘best balance’. Is there thus a tension between advocating the methodology of reflective equilibrium while urging a preference for the realist conception of theories (at least when these views are combined with an optimism rather than a scepticism about the possibility of success in theorising)? Certainly the methodology is compatible with the modelling

approach. And this combination of views could reduce the need for the tricky task of adjudicating between two good competitors which preserve different intuitions, since they could both be declared successful models (and perhaps appropriate for different purposes). It might then be suggested that the reflective equilibrium process can only provide a way of choosing theories on pragmatic grounds or according to subjective preference, and that it thus cannot be a route to the underlying truth behind the phenomena. So, the argument could continue, if we advocate that methodology we should also believe that there is no fact of the matter about one theory being objectively correct and we should thus take an anti-realist line. But there is nothing in the nature of the method of reflective equilibrium that commits us to such anti-realism. Compare the situation with inference to the best explanation. Someone might similarly argue that deciding on the best explanation can at best be a decision made on pragmatic and subjective grounds and cannot be guaranteed to deliver the correct explanation: there is no fact of the matter that one explanation is uniquely right. But such an anti-realism is not forced upon advocates of inference to the best explanation either: they can see their method as the way to reach the objectively and uniquely correct explanation.

If it turned out that there were ties for the best theory, as judged by §1 standards – two or more theories that were both sufficiently good to be judged successful theories of vagueness – then I might be obliged to rethink my commitment to reflective equilibrium plus realism plus a lack of scepticism. But we will be lucky if there is one adequate theory, let alone more than one. And if there is an outright winner, as I maintain, my combination of views is reasonable.<sup>3</sup>

In the next chapter I turn to the epistemic view – a theory offered as a realist account and defended as uniquely correct.

<sup>3</sup> There is, perhaps, an analogy between my response here and an attitude taken by Lewis (1994, p. 479). He considers the objection to his best-system account of laws that the standards of simplicity and strength (by which the best system is decided) are psychological, opening up the position to an undesirable idealist claim that what laws there are is up to us. He replies that if 'nature is kind' – a reasonable hope, he suggests – the best system will be robustly best, coming out first under any standards of simplicity and strength and balance. If nature is unkind, we would need to choose between a number of different (less appealing) responses; but he recommends that 'we not cross these bridges unless we come to them'.