Higher-Order Vagueness

Emma Bolton

Department of Philosophy
University of Sheffield

A thesis submitted in partial fulfilment of the requirements for the degree of Doctor of Philosophy

September 2018

Word count: 63,000

Abstract

This thesis deals with three interconnected questions: What does it mean to say that an expression is vague? Could there be 'higher-order' vagueness? And could there be a correct theory of the meaning of vague terms?

I argue that a traditional answer to the first question — that if an expression is vague, it has borderline cases and does not draw any knowable sharp boundaries — is incorrect. If this answer were correct, vague terms would exhibit higher-order vagueness. But I demonstrate a number of problems with this implication. I then put forward an alternative characterisation of vagueness: roughly, if a term is vague, speakers will fail to reach a certain kind of consensus on how that term is to be applied.

In answering the second question, I discuss some paradoxes that seem to emerge from the existence of higher-order vagueness. These paradoxes present a serious threat to the possibility of making sense of higher-order vagueness. I show that one such paradox can be generated using fewer resources than has previously been suggested, and so that this threat is greater than it may have seemed. I then argue that an attempt to avoid these paradoxes, Bobzien's 'columnar' theory of higher-order vagueness, fails to provide a plausible, or consistent, picture of higher-order vagueness.

Finally, I make use of my alternative characterisation of vagueness to argue that there is no correct semantics for vague terms. I discuss some ways in which theorists about vagueness could make sense of this, reflecting on the aims that we might have when approaching logic and semantics in the first place.

Acknowledgements

First and foremost I am grateful to Rosanna Keefe for her help in writing this thesis — for her engaging discussion and comments on really everything I've written over the past four years, for her encouragement, and for always pushing for more clarity when talking about the unclear. My two secondary supervisors, Jess Leech and Eric Olson, were also invaluable to the process: Jess in particular in helping me to develop a writing style, and Eric in helping to make this thesis make more sense 'from the outside'.

From the philosophical community in Sheffield (and its honorary members) I'd like to extend thanks to Simon Barker, Graham Bex-Priestly, Josh Black, Andrea Blomqvist, Lewis Brooks, Wungtei Buchem, Giulia Casini, Charlie Crerar, Matt Cull, Anton Eriksson, Ahmad Fattah, Trystan Goetze, Will Hornett, Miklos Kurthy, Damiano La Manna, James Lewis, Eleanor Marinho, Neri Marsili, Jed Martin, Josh Matthews, Josie McQuillan, Nadia Mehdi, Luisa Meytre, Robbie Morgan, Ashley Pennington, Cristina Roadevin, David Strohmaier, Rosa Vince, Sabina Wantoch, and Karsten Witt. All in their own ways they've brought joy to the time I spent writing this, and doing philosophy with them has really embodied what I think doing philosophy should be about. On this score, I'm especially thankful to Corin Bruce and Michael Greer, for the late nights we each spent talking, and for introducing me to the most important ideas I encountered while writing this.

For their friendship, support, and encouragement (and for some impossibly long email threads!), I'd like to thank Heli Himanka, Michael Krawec, LM, Cassi Roberts, Zaki Sidri, and Jess Walker — I couldn't have done this without them. I'd also of course like to thank my whole family for all the ways they've supported me.

The White Rose College of the Arts and Humanities funded my PhD — thanks obviously to them. And thanks to all those who I met at the Munich Centre for Mathematical Philosophy for a great stay — I learned a lot. Thanks also to the audiences of the seminars and conferences I presented at, in particular: numerous graduate seminars in Sheffield, the Joint Session in Cardiff, the 'Bounds of Logic Reloaded' conference in Moscow, the Work in Progress seminar in Munich, and the 'Pataphysique Sans Frontières' conference at the Facoltà di Tetrapiloctomia in Turin.

Finally, I would like to thank *you*, the reader. By reading this work, you're making its creation more worthwhile. Or are you just here to see if you're in the acknowledgements?

Contents

Introduction

Imagine some grains of sand that you'd be happy to say made a heap. Could taking one grain away from them stop them from making a heap? (It seems) no; so that same collection of grains but with one grain fewer must be a heap, too. Indeed, it seems that taking a grain away from *any* heap couldn't turn it into a non-heap. But then, would we still have a heap if we one by one removed all of the grains but one? It seems not. So how should we make sense of all of this?

Vagueness attracted me as a topic because it presents a set of problems (at the heart of them 'sorites' paradoxes of the kind just described) which give the appearance of being unsolvable — the answers you can give have a habit of crumbling in your hands. Learning about the 'standard' theories of vagueness, all of them seemed (to me, at least) to miss something important (or to fail to satisfactorily address it). In a way, the goal of this thesis is to articulate what exactly this 'something' is, and how to properly fit it into a theory of vagueness. One name that we could give to this 'something' is 'higher-order vagueness'. In informal terms, and in its broadest sense, higher-order vagueness is vagueness about where the effects of vagueness start and stop. To give an example of higher-order vagueness in action, let's say that we've got a theory that tries to solve the sorites paradox by saying that, while one grain of sand isn't a heap, and 10,000,000 might be, between these two extremes there are some cases which are in some way indeterminate: it's neither true nor false to say of them that they're heaps. That way there's no grain of sand that can be truthfully said to make a non-heap into a heap. We can illustrate what a *higher-order vagueness* problem for this solution to the sorites paradox looks like by asking 'Is there a grain of sand that takes a determinate heap to something that isn't *determinately* a heap?'. This leads to its own form of the sorites paradox, this time around 'determinate heap', rather than 'heap'. And this is by no means all there is to higher-order vagueness, as we'll see.

Higher-order vagueness problems (such as the new sorites paradox just described) seem to iterate indefinitely — even if you can solve one version of them, in doing so you will generate a new problem. In this way, higher-order vagueness poses a challenge to any theory of vagueness: don't just tell us how to make sense of the apparent lack of sharp demarcation of heap from non-heap, nor the apparent lack of sharp demarcation of determinate heap from indeterminate, but tell us how to make sense of the apparent lack of significant sharp demarcations *altogether* when a term like 'heap' is vague. This thesis is about what happens when you take this challenge seriously.

As it turns out, the implications are quite far-reaching. For one, we'll see that higher-order vagueness generates tension with other 'common sense' things that we might think that vagueness involves, in particular with the idea that a term's vagueness involves its having borderline cases, and the idea that it's not possible to know (say) that some particular grain of sand could make the difference to whether an object is a heap or not. This will be a reason to re-evaluate what vagueness even *is* and to question these common sense ideas. In turn, we'll see that this re-evaluation has some greater consequences — the last two chapters of this thesis deal with the idea that there could be no correct theory of the meanings of vague terms whatsoever. This is a significant result in itself, but it implies more broadly that there are certain limitations on theorising about natural languages (such as English) in general, and perhaps motivates some different ways of doing this theorising in the first place — this case study should serve as a cause to re-evaluate our goals when investigating logic and semantics.

Here's a broad chapter-by-chapter overview. Chapter 1 is about orthodox ways of saying what vagueness actually consists in to begin with, identifying three interconnected features that vague terms are often taken to have. It's said that vague terms don't draw any knowable 'sharp boundaries' (in the sense, for example, that it's not possible to know that some particular grain of sand could make the difference between a heap and a non-heap), that they have borderline cases (so cases about which it's not possible to know whether it's a heap or not), and that they are higher-order vague.

The two chapters which follow this work through some difficult consequences of taking vague terms to have these features. Chapter 2 deals with a set of challenges to the very idea of a certain kind of higher-order vagueness. A number of writers have presented a variety of 'paradoxes of higher-order vagueness', each attempting to show that the existence of higher-order vagueness would be in itself contradictory. In this chapter we'll see how a selection of these paradoxes work, and I'll present my own paradox of higher-order vagueness, which I take to rely on the fewest assumptions of all of them.

Chapter 3 looks at what could happen if even *these* few assumptions were abandoned. The result of this abandonment is what we can call a 'columnar' view of higher-order vagueness — so called because they divide up objects to form three monolithic categories which look like columns when represented visually. Susanne Bobzien has offered a sustained defence of this kind of view, so chapter 3 focuses on her view in particular. It offers an explanation of how her version of this kind of view works, while offering two key criticisms. This serves to expose some of the flaws of columnar views in general, and so this chapter will show why we should rule out these views in general.

Chapters 2 and 3 will have cast serious doubt on the possibility of understanding higher-order vagueness (in the sense in which it's construed in those chapters) in a way that's both consistent and plausible. But since higher-order vagueness and the two other 'orthodox' features of vagueness that I mentioned when talking about chapter 1 are all tightly connected, this result casts doubt on those other two features being features of vague terms. As such, chapter 4 aims to arrive at a new way of saying what vagueness is which isn't so closely tied to higher-order vagueness (again, in the sense in which it's been dealt with that far — there's more than one way we can understand higher-order vagueness, as we'll see). We'll look at two proposed answers to this 'characterisation' question from Eklund and Horgan respectively, and see how they fare against a set of test cases. In the end we'll arrive at a new answer, which bears more resemblance to Horgan's than Eklund's, and I'll show how it handles the test cases. Put (very!) roughly, I'll suggest that a necessary condition for an expression's vagueness is that competent speakers won't necessarily agree on when to apply that expression to different objects, when to say it doesn't apply, and when to give some other response altogether, when considering gradual changes to those objects (such as adding or removing grains of sand from heaps). This is only a necessary condition, but it's all we need for now, and I'll gesture towards a way of reaching a set of necessary and sufficient conditions.

Chapter 5 makes use of the necessary condition for vagueness from chapter 4 to argue for probably the central claim of this thesis, that there could be no correct semantic theory for vague terms. It does this by arguing that, first, given that vague terms all satisfy this necessary condition, a correct *semantics* for those terms would need to satisfy a corresponding condition, namely that it didn't categorise all objects into those to which the relevant term applies, those to which it doesn't, and any others. From here, the chapter looks at a set of arguments that make use of this claim about the semantics of vague terms (or similar claims) that try to show that vagueness is in some sense incoherent (or 'impossible'), in particular focusing on an argument from Fine, but also ones from Sainsbury and Horgan. We'll see how Fine's argument can be made to work with even fewer resources than Fine suggests, while nevertheless claiming that it's not completely convincing given its reliance on the 'reductio' rule of inference. After that, I'll present an argument along related, but quite different, lines, that aims to show that there could be no correct semantics for vagueness.

The final chapter tries to deal with some of the 'fallout' of the results established in chapter 5. It tries to show that, while the result *does* show that no theory of vagueness could strictly speaking be correct, any of these could nevertheless be good (or even the best) theories of vagueness. It offers some ways of understanding how this could be so, looking at different approaches to why we study semantics in the first place — borrowing a taxonomy of approaches from Shapiro — and seeing how they can be made to fit with the claim that there's no correct semantics for vague terms. The chapter then concludes by outlining some ways in which the impossibility of a correct semantics could actually help theories of vagueness in solving traditional problems of vagueness, and the extent to which higher-order vagueness can be made to fit into those theories.

Chapter 1

Characterising Vagueness and Higher-order Vagueness

This chapter will set up some of the key topics of the thesis. First, I'll look at some ways in which we could characterise vagueness. This is something of an important task if we want to start theorising about vagueness: before we start, we ought to get clear on what we're theorising about. I'll provisionally settle on a view made explicit by Greenough, and at least implicitly endorsed by others, that vague expressions minimally (that is, characteristically) exhibit what we can call 'epistemic tolerance' and 'epistemic' borderline cases — roughly stated, that small changes can't make a *known* difference to whether a vague expression applies or not, and that there are some cases about which it's not possible to tell whether that expression applies to them. We'll revisit the question of how to characterise vagueness much later, but for the time being we'll settle on this characterisation in order to investigate it.

The second thing to do is to show how this characterisation of vagueness entails that vague expressions exhibit 'higher-order' vagueness. There are (at least!) two things that we could mean by 'higher-order vagueness'. I'll explain these two particular understandings of the term and show why vague expressions are higher-order vague in both senses if they are characteristically both epistemically tolerant and have epistemic borderline cases.

By this point we'll have some strong reasons to think that vague expressions are higher-order vague, then. This will leave open a natural set of questions about the nature of higher-order vagueness. At this point I'll introduce some formal apparatus to help us address these questions: sentence operators 'B' and 'D', representing the expressions 'borderline' and 'definite', which we can understand as analogues of certain operators from modal logic. From there, we can understand certain questions about higher-order vagueness as questions about the logic governing these two operators, and so we'll have effectively set up one of the broad topics of the thesis.

1 Characterising Vagueness

1.1 Setting Up the Question

'Old', 'blue', 'tall', and 'trousers' are plausibly vague terms. What unites them and makes them *vague*? In one way this is a difficult question to answer, because 'vague' has a very particular meaning in the philosophical literature on vagueness. Shouldn't this make the question of what vagueness is easy to answer? Unfortunately, no — there's no 'canonical' statement of what vagueness is. But let's think about how we can get a grasp on the phenomenon underlying the expression 'vague' as it's used within philosophical discussions of vagueness. One way is to think about the sorites paradox — this is probably the first place where vagueness is addressed in Western philosophy, even if not under that name. Forms of the sorites paradox go roughly along these lines (though see for example Hyde 2011 for more on the history of this and some of the various forms it can take): Take a collection of grains of sand that you think is big enough to be a heap. Now, *surely*, removing a grain of sand from this heap couldn't make it into a non-heap. So a collection of grains of sand with one fewer grain is a heap, too. And, likewise, one grain fewer than *that* must be a heap. We can keep going with this reasoning until we're forced to say that a single grain could be a heap, and at this point something seems to have gone wrong.

We can construct analogous paradoxes for other vague expressions: making people successively older for 'young', or adding hems to trousers until we get shorts. A few kinds of claim are involved in generating all of these instances of the paradox, which look plausible enough by themselves. One is that vague expressions apply (or could apply) in some cases and don't (or wouldn't) apply in others: there are trousers, and then there are shorts. Another is that small changes of a relevant kind appear insufficient for us to say that a vague expression applies before that change, yet fails to apply after it: could there really be a shortest tall person and a tallest non-tall person, separated by a fraction of a millimetre? Another is that we can use these first two facts to create a chain of reasoning showing that a sufficiently *big* change of the relevant sort is, after all, insufficient for a vague expression to apply before that change, yet fail to apply after it.

Another way to think about it. Consider how you might make judgements about the actual cases we're considering when we think about the sorites paradox. We can start with cases we're confident about: Violet Brown (the oldest woman in the world) is old; a cloudless midday sky is blue. But we soon find ourselves confronted with cases where we hesitate: I don't think of my parents as old, but perhaps they are. And what colour are the walls of the Information Commons in Sheffield? To put this another way, it seems that vague expressions have clear cases and borderline cases.

What I've said here points to three broad phenomena that we can associate with vagueness.[1] The first is that vague expressions appear not to draw any sharp bound-

[1] These are identified by, for example, Keefe (2000, 7-8), Greenough (2003), and Smith (2008, 1-2).

aries. The second is that they give rise to the sorites paradox. And the third is that they have borderline cases.

Could any, or some combination, of these be the characteristic feature or features of vagueness? In giving an answer to this question, one thing to bear in mind is that the question of what characterises vagueness is in some sense 'prior' to actually theorising about it: we should really know what we're theorising about before we start developing the theory. But, of course, there are already theories of vagueness out there. So, in looking for characteristic features, we must be careful not to tread on the toes of the people those theories belong to. If we arrive at a characterisation of vagueness which actually rules out some particular theory, we'll effectively need to say that that theory just *misunderstands* the phenomenon. This isn't something we'll want to say lightly, so at this stage we need to maintain a certain degree of theory-neutrality, and be willing to violate it only if we have very good reasons. With this in mind, let's try to make the 'minimal' claims we could make about vagueness more precise, and see if we can take any of them to characterise vagueness.

1.2 No (Knowable?) Sharp Boundaries and Tolerance

We can start by looking at the idea that it might be characteristic of vague expressions that they do not draw any sharp boundaries. This is the idea that vague expressions 'tolerate' small changes: if a vague expression applies to an object, it still applies after a sufficiently small change of the relevant sort. Following Wright (1975, 334), we can say that an expression is tolerant when small changes of these sorts are unable to affect whether or not it applies.[2]

[2]Wright gives his notion of tolerance as follows: 'F is tolerant with respect to [a concept Θ, such that a case to which F applies can be transformed into a case where F does not apply by a sufficient change in respect of Θ,] if there is also some positive degree of change in respect of Θ insufficient ever to affect the justice with which F is applied to a particular case' (1975, 334). I'm not concerned with whether or not the notion of tolerance that I'm making use of in this chapter matches up with Wright's exactly.

This is certainly a tempting thought. Isn't it *just obvious* that expressions like 'green' and 'old' couldn't suddenly start applying once you add the smallest amount of blue paint to that mixture that's more on the yellow side, or the moment you blow out the candles on that crucial birthday cake? But, while we might make it a priority that any theory of vagueness we subscribe to must ensure that vague expressions are tolerant, we shouldn't automatically rule out theories which, for example, take vague expressions to draw (albeit unknowable) sharp boundaries (see for example Williamson (1994)). It might sound like a mistake, but if we want to reject this kind of theory, we'll need to show why it's a bad explanation of the phenomenon (if it is!) rather than an attempt to explain an altogether different phenomenon.

To be a bit more conciliatory towards such theories, here's another minimal characteristic that we might take vague expressions to have: that they *seem* to draw no sharp boundaries. Such a broad statement of this possible characterisation isn't going to be good enough without some more refinement. For one, to whom does it have to seem this way for an expression to be vague? And what counts as appearing to draw no sharp boundaries?

Greenough gives us a way to sharpen this idea by introducing the notion of 'epistemic tolerance'. An expression is epistemically tolerant when a small change of some particular kind cannot make a known difference to whether that expression applies or not (2003, 258). The thought is that it's just not possible to know, for example, which the heap with the fewest grains is, or how long the shortest trousers are. We can make use of this notion to make precise the idea of vague expressions apparently lacking sharp boundaries: they appear to draw no sharp boundaries, the thought would go, in that we'd be unable to detect any if we tried, even if we were able to investigate in any way we wanted to.

What is the significance of this last part: being able to investigate in any way we wanted to? Without this element of epistemic tolerance, we might worry that some plausibly non-vague expressions might turn out to be vague for the wrong reasons. For example, presumably it's not possible to tell, just by looking, the day on which a growing oak tree measures more than 30ft. Yet 'more than 30ft' seems to be precise, or at least isn't vague because of this human limitation on detecting the heights of trees by eye. So we'd need to specify that epistemic tolerance is a *specific kind* of limitation on knowledge: while we can measure an oak tree on a given day to find out its height, no amount of measuring will tell us the first day on which it is tall. In other words, it needs to be impossible to tell where a vague expression starts and stops applying, in a way which resists enquiry and new information.

We can illustrate this notion of 'enquiry resistance' a little more here, though I don't intend to settle the question of what counts as resisting enquiry. Sorensen (instructively) likens questions such as 'where do you draw the line?' (in the case of vagueness) to questions such as 'does a glass house have any windows?' — even a persistent enquirer 'hits a wall' in trying to answer them (2001, 24).[3] Although it's possible to give answers to these questions, it seems to just be a feature of the questions themselves that we couldn't *settle* them, and it likewise seems that no reasons that we could point to in answering them could really be conclusive. And, as Sorensen puts it, even if there were a reliable system for answering these questions correctly (assuming there were such correct answers), 'if Mr Vagueluck happens to adopt the system, his reliably correct beliefs about the truth-values of [sentences claiming a line is drawn] would not constitute knowledge. For his beliefs would have both a lucky origin and be sustained by luck' (2001, 25).

So far, then, we have an answer to the question 'What should count as appearing to draw no sharp boundaries?' — it's when it's not possible to tell where a vague expression starts or stops applying in a way which resists enquiry (that is, when it

[3]In this passage and the one that follows, Sorensen is actually talking about borderline cases, rather than counter-examples to tolerance, but his remarks apply just as much to the latter, and are also a trivial consequence of his epistemicist theory more generally.

is epistemically tolerant). But who does this need to be apparent to in order for an expression to be vague? Thinking in terms of epistemic tolerance gives us a fairly natural answer to this question: if an expression is vague, it's impossible for *anyone* to know that some small change makes the difference to whether it applies or not.[4]

It looks like epistemic tolerance is a feature which we can take (at least!) all vague expressions to have, then, even if we can't go as far as saying that they are tolerant (in the non-epistemic sense) straight away. Let's now look at the other characteristic features I mentioned before, to see if any additional features are essential to vague expressions.

1.3 Sorites Susceptibility

What about the second potential characteristic phenomenon of vagueness, that vague expressions give rise to the sorites paradox? In trying to make this more precise, we face some similar questions the the ones we addressed in the last section. What counts as giving rise to the sorites paradox? Does an expression need to *appear* to give rise to the paradox in order to be vague? If so, to whom?

Let's start with a fairly crude suggestion: if an expression is vague, we can make a 'sorites argument' using it. By this I mean that we can construct an argument that is similar (in some specific way) to this:

(1) 1,000,000 grains of sand are a heap.

(2) If n grains of sand are a heap, $n - 1$ grains of sand are a heap.

So (C) 1 grain of sand is a heap.

To draw out what counts as an argument being 'similar' to this argument in the relevant way, consider what happens if you replace 'grains of sand are a heap' with 'is a number greater than 10'. Clearly 'number greater than 10' is not vague — and clearly

[4]Although to avoid this lapsing into non-epistemic tolerance, we should make an exception here for omniscient beings, for example.

the argument is a bad one — so there's something that distinguishes the argument using 'number greater than 10' from the one using 'heap', and *that's* what we're interested in if we want a characterisation of vagueness in terms of the sorites paradox. What might this be?

We might say that a sorites argument using 'number greater than 10' is just obviously not a good argument, while it's at least not obvious where the sorites argument using 'heap' goes wrong, if it goes wrong at all. From there, we could proceed in a few directions. One bold strategy would be to say that the sorites argument using 'heap' is not just not obviously a bad argument, but is in fact a good argument; its premises are true and its conclusion follows (and so its conclusion is true as well). But this isn't going to get us far, since it rules out a large number of the theories of vagueness out there which try to *solve* the paradox by showing where sorites arguments go wrong. And anyway, if we go this way, we'll have to say that (2) is true, which we already rejected a form of as a characteristic claim about vagueness.

Maybe we'll do better if we link vagueness to the sorites paradox by saying that sorites arguments involving vague terms in some way *come across* differently to those with precise terms. For one, perhaps we could remain fairly close to the previous suggestion by saying that when we're dealing with vague expressions, sorites arguments seem convincing where they wouldn't if we were looking at precise expressions. But that's obviously wrong. I suspect that even philosophers who think the sorites argument is sound think so reluctantly; surely anyone's gut instinct on sorites arguments (if they're interested at all) is that *something* is wrong with the reasoning as a whole.

Would *that* work as a characteristic feature of vagueness? The idea would be that sorites arguments with precise terms appear to be obviously bad arguments, while their counterparts with vague terms appear 'paradoxical': they appear to have true premises, yet an apparently obvious falsehood appears to follow from them. Here's a way to illustrate why I worry about this suggestion. Consider a distant possible future in which humankind has been ravaged by nuclear war. A few generations pass and some English-speaking survivors come across fragments of Tim Williamson's writings on vagueness, stating that vague terms *do* draw (unknowable) sharp boundaries, and interpret them as a holy text. As a result, everyone learns from an early age that all expressions draw sharp boundaries, although they can't always tell where they are. But besides that, suppose, they use vague terms in much the same way as we do. Now, in this future, sorites arguments don't even look paradoxical: everyone 'knows' that removing a grain of sand *can* destroy some heaps. So, as it turns out, on the proposal we're looking at, there would be no vagueness in these people's language. Yet it seems that they wouldn't really have eliminated vagueness — their language just doesn't *appear* vague to them — and so it likewise seems that we should reject a characterisation of vagueness stated in terms of sorites arguments appearing paradoxical.

Characterising vagueness in terms of the sorites paradox itself, then, doesn't seem to work, so we'll need to go on to consider some other options for our minimal characterisation for vagueness. This is not, of course, to say that theories of vagueness shouldn't address the paradox as a pressing or central issue. It's just that we don't need to make reference to it in our minimal characterisation. Instead, we can note that the paradox is tied up with other questions about vagueness. In particular, insofar as a theory of vagueness addresses the question 'Does the fact that vague expressions draw no knowable boundaries indicate that they draw no boundaries at all?', it will need to say something about the sorites-shaped consequences that come along with it. To spell this out: if the answer is 'yes', are the premises of sorites arguments such as 'If n grains are a heap, $n-1$ grains are a heap' true? If so, sorites paradoxes loom. On the other hand, if vague expressions *are* taken to draw sharp boundaries, albeit unknowable ones, we'll have an explanation for why sorites arguments are in fact unsuccessful. Let's now go on to consider another possible feature that might characterise vague terms.

1.4 Borderline Cases

Do all vague expressions have borderline cases? We should really ask whether all vague expressions have *possible* borderline cases: even if borderline cases are characteristic of vagueness, presumably we couldn't make 'blue' precise by painting everything red, because it would still be possible for something to be a borderline case of 'blue'.

One thing to say in defence of this suggestion is that we can readily call to mind cases in which we've expressed some hesitation or uncertainty in applying vague expressions: there are surely people whose baldness is questionable, and bread that's on the iffy side of stale. Perhaps that's sufficient reason to think that any vague expression must have borderline cases.

This is a little quick. The idea of a 'borderline case' isn't easy to explicate, but we should get a little clearer on it before trying to say whether borderline cases are characteristic of vagueness. That said, I'm only going to do this in broad strokes here.

Two notions of 'borderline case' which we can presumably rule out as especially relevant to vagueness are pointed out by Susanne Bobzien (2013, 5): we sometimes say that something is a borderline case when the relevant expression *only just* applies, and we can likewise say that something is a borderline case if the relevant expression *doesn't quite* apply.[5] For example, we might sensibly say that St Davids is a borderline case of 'city' in that it only just qualifies as one, and that whales are borderline fish in the sense that 'fish' only just fails to apply to them.

Leaving those aside, we can introduce two more notions of 'borderline cases': we can talk about semantic borderline cases and epistemic ones. We'll look at these in turn. I'll take semantic borderline cases to represent some 'middle ground' between truth and falsity — this could mean that a sentence such as 'The walls of the IC are

[5] This distinction is also brought out under different terminology in Raffman (2005).

blue' comes out as neither true nor false, as having some 'indeterminate' truth-value, as being both true and false, or true to some degree between 0 and 1, to give some examples.

Could semantic borderline cases be characteristic of vagueness? It seems not, for reasons which are perhaps predictable now. Not all theorists about vagueness are going to agree that there *are* semantic borderline cases. For example, epistemicists such as Williamson will say that while we perhaps cannot know whether 'lengthiest' is a long word or not, it's uniquely either true or false to say that it is. And we shouldn't rule out this theory as not really being a theory of vagueness.

On the other hand, the epistemicist's treatment of borderline cases points nicely to a more neutral 'epistemic' sense of 'borderline case'. Broadly construed, it is not possible to find out whether an expression applies to an *epistemic* borderline case of that expression, even if we have all of the relevant evidence and are in the right kind of conditions to find out. To simplify things for now (we'll see how things can get more complicated later!), let's focus on *knowledge* as the epistemic state that the 'epistemic' aspect of this kind of borderline case relates to. So, for example, darts is (probably) an epistemic borderline case of 'sport' in that it is (probably) not possible to know whether or not it is a sport.

Now, at the beginning of this section I gave some reasons I hoped would lend some initial plausibility to the thought that vague expressions have epistemic borderline cases. Greenough goes one further in giving us a proof, showing that if we're already convinced that epistemic tolerance is characteristic of vagueness, we should also think that epistemic borderline cases are. Informally, the reasoning goes as follows (2003, 267-9).

Suppose that an expression, say, 'tall', is epistemically tolerant, and yet suppose that it has no (epistemic) borderline cases. In having no borderline cases, every object to which 'tall' could apply can be known either to apply or not. Yet since 'tall' is epistemically tolerant, if it's possible to know that some object is tall, it's not possible to know that some object sufficiently similar in height is not tall: this would amount to a known change from tall to not-tall on the basis of only a small change.

So far, so good. Now, we know that someone who's 8ft (say) is tall, and someone who's 5ft tall is not. We can start the crucial chain of reasoning from here. Take away any incredibly small amount of height from the person who's 8ft tall. Then, because 'tall' is epistemically tolerant, it's not possible to know that the resulting person is not-tall. And yet we assumed before that, for every person, it's possible to know either that they're tall or that they aren't. So, since we don't know that this slightly shorter person *isn't* tall, we must know that they *are* tall. And we can keep repeating this reasoning until we conclude that we know that the 5ft person is tall, contradicting the plausible assumption that they are not. So assuming that 'tall' is epistemically tolerant but has no borderline cases leads to a contradiction, and so we seem to have shown that if an expression is epistemically tolerant, it also has borderline cases.

The thought that's ultimately driving this is pretty compelling, that if vague expressions don't draw any knowable sharp boundaries, there must be some objects which act as a kind of 'buffer zone' — the borderline cases. Note, though, that Greenough's argument relies indispensably on at least some resources from classical logic. In particular, the conclusion only follows if we're allowed to use reductio ad absurdum, and then (an implicit instance of) double negation elimination, at the end. Still, let's grant this: whether this argument works or not, in order to see the consequences of the claim that vague expressions have borderline cases, we'll need to assume that that claim is correct.

2 Higher-order Vagueness

Let's briefly recap. So far we're assuming that all vague expressions are epistemically tolerant — they draw no knowable sharp boundaries — and exhibit borderline cases — cases about which it is not possible to know whether the relevant expression applies. We'll now see that these assumptions lead to a perhaps surprising consequence, that what we can call 'higher-order vagueness' exists in two key senses.

Actually, first I should point out that there's a third sense of 'higher-order vagueness' that I'm going to put to one side without a great deal of comment. This is the sense which takes the first 'order' or 'level' of vagueness to be the vagueness of expressions such as 'old', 'tall', and all the familiar ones, and the second 'order' of vagueness to be the vagueness of the term 'vague' itself. Is 'vague' vague? Given our assumed minimal conditions for vagueness, this question amounts to asking whether there's a knowable sharp boundary between vague and non-vague expressions, and whether there are borderline vague expressions. These are interesting questions, but it would be too much of a diversion to directly address them here (though see Sorensen (1985), Hyde (1994) and Varzi (2003) for discussion).

2.1 The Vagueness of Apparatus Describing Vagueness

On to the two 'key' senses I mentioned before, starting with what I take to be the more general of the two. So far, our minimal characterisation of vagueness makes a lot of reference to knowledge: epistemic tolerance amounted to an expression's having no knowable sharp boundaries, and borderline cases (minimally) were cases where we can't know whether the relevant expression applies. Now, is 'know' vague? Perhaps that's not quite the right question. We know that, say, we can't know where the sharp boundary between trousers and non-trousers lies — which hem makes the difference — but is it possible (and here's where things start getting complicated) to know of some garment that it's possible to know that it's trousers, and yet know that, if it were hemmed once, it wouldn't be possible to know that it was trousers? To put this another way: is it possible to know where the line is between garments you can know to be trousers, and the borderline cases of 'trousers'? Presumably the answer is 'no', as it's going to be in no way clear at what point justification — in the belief that some garment is trousers — drops off over successive small changes in spite of willingness to believe; it's likewise going to be difficult to pinpoint in all cases the point at which someone's hesitant belief turns to full-on cold feet.

This kind of idea iterates. I've been holding off from introducing any formal apparatus this far, but at this point we're coming close to needing it. Let's try without, for now, and perhaps what follows will provide some motivation for approaching higher-order vagueness 'formally' given how ugly some of our constructions will get without doing so. So, we have some garments which we can know to be trousers, some of which we can know we can know to be trousers, *some of which* we can know we can know we can know to be trousers. Take one of these last kind of garments, and think about successive hems we could make to it. Is there a point at which we can know that we can no longer know that we can know that the garment in question is a pair of trousers? Probably this question seems a little trickier to answer, owing to its quite complicated structure. But what small change could we possibly know to draw this line? Could it really become a precise matter just by moving a 'level' up? Indeed, could moving up any number of levels make this precise? It seems not.

At this point it seems worth pointing out that this phenomenon isn't related exclusively to the vagueness of 'possible to know' or 'epistemic borderline case'. We can see similar considerations coming out if we consider a proto-theory we might start developing on the basis of our minimal characterisation of vagueness. Such a theory might say that we can't know whether 'tall' applies to borderline cases because there's nothing to know — it's neither true nor false that a given borderline case is tall. Suppose this theory is getting things broadly correct. Then, in shrinking an uncontroversially tall person (André the Giant, say), could we identify a millimetre that takes it from true to say they're tall to it being neither true nor false that they're tall? It strikes me as unlikely. And this will iterate just as before. To give just one iteration, there likewise seems to be no identifiable point where it stops being true that it's true to say that someone is tall, and becomes neither true nor false that it's true to say they're tall.

The general moral here is this: given our characterisation of vagueness, we're not going to be able to describe vagueness fully using exclusively precise terms. Vagueness won't turn out to ultimately be a kind of precision in that no theory which is really a theory of vagueness should be able to give a precise description, case by case, of how we get from positive cases of an expression to the negative cases over a succession of small changes. This is our first sense of 'higher-order vagueness'.

To see why vague expressions characteristically exhibit this kind of higher-order vagueness, consider a predicate artificially introduced by Sainsbury (1991, 173), 'child*'. Someone is a child* if they're 16 or younger, is not a child* at 18 or over, and for anyone else it's indeterminate whether they're a child*. Is 'child*' vague? It seems not. All the cases are covered by the definition. Yet it's not possible to know that a sixteen-year-old is a child* and a seventeen-year-old is not — and so 'child*' is by definition epistemically tolerant — and likewise seventeen-year-olds are (epistemic) borderline cases of 'child*'. Our minimal characterisation was therefore missing something, and we're now in a position to see what: 'child*' isn't vague because we can apply our vocabulary describing vagueness to 'child*' in a precise way. While 'child*' is epistemically tolerant, we can identify the exact second which takes us from a sixteen-year-old who we can know that we can know to be a child* to a seventeen-year-old of whom we can know *that we don't know* whether or not they're a child*. Likewise we get a sharp transition from determinately determinate children* to determinately indeterminate cases, even if there's no sharp transition from determinate children* to determinate non-children*. For these reasons, 'child*' is ultimately precise, and, more generally, expressions need to be higher-order vague in order to be vague.

2.2 Higher-order Borderline Vagueness

Now let's go from the general to the specific. It's a little easier to spell out what I mean by the second sense of 'higher-order vagueness' that I want to introduce here, and it's worth noting here as well that this is more often the sense of 'higher-order vagueness' that people writing about vagueness have in mind. We saw just now that expressions like 'borderline case', as they apply to vague expressions, are themselves vague. The second sense of 'higher-order vagueness' concerns the consequences of this for the notion of borderline cases.

We know (suppose!) that 'old' is vague, and so has borderline cases. Consider the expression 'borderline old'. Some people are borderline old, and some people aren't. And we saw in the previous section that the difference between being borderline old and non-borderline old doesn't seem to be a precise matter: 'borderline old' is itself vague. Are there, then, borderline cases of 'borderline old'? If so, 'old' itself exhibits 'higher-order vagueness' in the second sense.

We should really call this kind of higher-order vagueness 'higher-order *borderline* vagueness', to distinguish it from the kind I mentioned before, but for the rest of this section I'll omit the 'borderline' bit for ease of reading. Here's a statement of what this sense of higher-order vagueness amounts to, then. First, call a borderline case of 'borderline old', say, a 'second-order' borderline case of 'old', and a borderline case of 'second-order borderline case of 'old'' a 'third-order' borderline case, and so on. We'll say that an expression is first-order vague when it has borderline cases, and more generally that it is nth-order vague when there are nth-order borderline cases of that expression. Accordingly, we can call an expression 'higher-order vague' when it's second-order vague or higher, and 'radically' higher-order vague when it's nth-order vague for all n.

Given the characterisation of vagueness we're currently working with, we get a fairly natural strategy for showing that all vague expressions are radically higher-order vague. Take 'old' again, for example. We know that there are borderline cases of 'old'. So it has first-order borderline cases, and we're off to a good start. Now, we also know that 'borderline case of 'old'' is itself vague, and so, since vague expressions have borderline cases, there are borderline cases of 'borderline case of 'old''. So we've got second-order borderline cases, too. We can keep going like this, exploiting the vagueness of 'borderline case of...' for each successive kind of borderline case, in sum showing that 'old' is radically higher-order vague.[6]

This argument certainly seems simple enough, but perhaps you noticed that I didn't give any examples of the higher-order borderline cases I purported to show the existence of with this argument. To be honest, I lose the intuition that any object is a borderline case at 'orders' above two, and so I find it difficult to give definitive examples of such cases. I wonder if perhaps I'm borderline tall (at 6ft) relative to, say, adult males in the UK, so perhaps I'm a borderline borderline case of 'tall'. But it's much harder to imagine saying that someone was a millionth-order borderline case of 'tall'. Still, let's suppose that the argument works.

2.3 A Task for Theories of Vagueness

So far, then, we're working on the assumption that vague expressions are epistemically tolerant, exhibit epistemic borderline cases, and are higher-order vague in both senses just discussed. Given all this, which questions might we expect a theory of vagueness to address? Some natural questions are: Does the fact that we can't identify any sharp

[6]This argument is gestured to in Greenough (2003, 277-8), for one.

boundaries when it comes to expressions like 'old' indicate that there are no such boundaries? And, what does the fact that we can't identify whether vague expressions apply to borderline cases or not tell us about the *semantic* status of borderline cases (if anything)? Does the fact that we can't identify this either way imply that it represents some gap between application and non-application of that expression, for example? Probably it goes without saying that these questions open something of a can of worms. But I'm more interested here in questions generated by higher-order vagueness. Here's the one that will occupy the next two chapters: Given that there are infinitely many different kinds of borderline case (first-order, second-order,...) associated with expressions like 'child', 'red', and so on, which could be exemplified over a series of small changes of the right sort, what are the structural relations between these different categories of borderline case?

This question perhaps doesn't come across as the most interesting question to focus on. But in getting at an answer, we'll see that we run into problems fairly quickly, and end up making some interesting, and often counterintuitive, claims. All this promises to reveal something fundamental to the phenomenon of vagueness.

3 The Logical Apparatus

3.1 A Digression: Modal Logic

The time has finally come to introduce some formal apparatus! We've reached a question predominantly about *structure*, and this question will be much easier to answer if we can make use of tools well-suited to making claims about structure. But first we'll need to take a small detour into questions about possibility and necessity to see just which tools we need.

When investigating structural questions about different senses of possibility and necessity, rich systems of logic have been developed under the broad umbrella of 'modal' logic. The basic language of systems of modal logic often makes use of (parts of) the language of propositional logic — connectives $\neg$, $\wedge$, $\rightarrow$, and so on, propositional variables $p_1, p_2, \ldots$, and metalinguistic variables $A_1, A_2, \ldots$ standing for arbitrary sentences in our language (I'll often just write 'A' as a metalinguistic variable when no confusion could result). We can also make use of the language of first-order logic, but for simplicity's sake we'll leave this to one side for now. The language will add to these a primitive sentence operator, either $\Diamond$ or $\Box$, to represent 'possibly' or 'necessarily' respectively. This allows us to form sentences '$\Diamond A$' or '$\Box A$', representing 'It's possible that A' and 'It's necessary that A' respectively. So '$\Diamond p_1$' will be a sentence saying 'It's possible that $p1$' and '$\Box(p_1 \rightarrow p_2)$' will be a sentence saying 'Necessarily, if p_1 then p_2', to give some examples. Whichever of '$\Diamond$' or '$\Box$' we start with, $\Diamond$ and $\Box$ are duals,[7] so we can subsequently introduce either by defining it in terms of the other in perfectly symmetrical ways: $\Diamond A$ will be equivalent to $\neg \Box \neg A$, and $\Box A$ will be equivalent to $\neg \Diamond \neg A$.

With this language in place, we can start to approach structural questions about necessity and possibility. One way to approach this is through *syntax* — as Hughes and Cresswell put it, we could specify a set of axiomatic claims in the language, and give a set of rules which tell us how we can manipulate the sentences in our language and go from one to another (1996, 23-4). All of this is intended to in some way reflect the structural properties of claims about possibility and necessity. To give some examples, we might take

(T) $\Box A \rightarrow A$

to be an axiom, and we might introduce a rule that if a sentence A is either an axiom or can be derived from an axiom (that is, if A is an axiom or a theorem), then $\Box A$ is

[7] An operator O is a dual of an operator P when OA is equivalent to $\neg P \neg A$.

too. We can take these in turn to represent that if a sentence is necessary, then it's true, and that if a sentence is provable, it's provable that it's necessarily true.

Another way to approach modal logic is through *semantics*. We can do this by looking at 'models', which offer interpretations of our basic language. We can make these quite complicated, but, for example, Chellas defines 'standard' models in systems of modal logic as structures $< W, R, P >$ — that is, ordered triples made up of elements W, R and P. W here is a set of 'worlds', R is a binary relation that can hold between these worlds, and P is a function from atomic sentences and worlds to truth-values. In a model, then, we have a set of worlds, some of which might be related to one another by the R-relation, with the function P determining which atomic sentences are true at each of these worlds. In turn, all this is enough to determine which non-atomic sentences are true, too (Chellas 1980, 67-8). For example, if p_1 and p_2 are true at a world, $p_1 \wedge p_2$ will be true at that world.

This last point gets a little more complicated when we start thinking about sentences containing '$\Box$' and '$\Diamond$' because, while we can derive the truth of $p_1 \wedge p_2$ from the truth of p_1 and the truth of p_2, it seems that knowing whether a sentence p_1 is true isn't really enough to know whether it's *necessarily* true. And likewise knowing that p_1 is false isn't enough to say whether p_1 *could* be true. In this sense, '$\Box$' and '$\Diamond$' are not *truth-functional* operators, and so to get the truth-value of a sentence '$\Diamond A$' or '$\Box A$' at a world, we need to look beyond the truth-value that P assigns to A at that world, and look at the worlds 'around' the particular world we're interested in, making use of the R-relation. The basic idea is this: a world α in a model considers any world it can 'see' to be a possibility, and so if α can 'see' at least one world in which a sentence A is true, then A is possible from the perspective of α. Likewise if A is true in every world that α can see, then A is necessarily true from the perspective of α. The R-relation can be thought of as just this 'seeing' relation, giving the following standard truth-conditions (as in for example Chellas 1980, 68): $\Box A$ is true at a world α in a model M if and only if every world β in M such that $R\alpha\beta$ is such that A is true at β. And $\Diamond A$ is true at a world α in a model M if and only if some world β in M such that $R\alpha\beta$ is such that A is true at β.

Evidently, this R-relation is quite important. If we adopt this semantic approach, one of our key questions will be: which worlds should we take the R-relation to hold between? To give a flavour for how we might answer this, here's an example. If we say that, in our models, every world can see itself, the truth-conditions for '$\Box$' will tell us that if a sentence is necessarily true at some world, it will be plain true in that world, too. For a lot of senses of 'must' or 'necessarily', it's therefore going to be desirable to say that the R-relation has this property of *reflexivity*.[8]

This last example should hopefully gesture, as well, at the fact that the semantic and syntactic approaches to modal logic can be seen as two sides of the same coin: endorsing

(T) $\Box A \to A$

from earlier does much the same work as saying that the R-relation between worlds is reflexive.

3.2 'B' and 'D'

How does this relate to the question of higher-order vagueness? Well, the first thing to note is that the kind of language we can use to talk about higher-order vagueness is very similar to the language of possibility and necessity. Consider that we can say of someone who's borderline old that they *might* be old or they *might not*, for example. We have this well-developed system for making claims about necessity and possibility

[8]Though note that even this is going to vary depending on which sense of 'must' we're interested in: 'All students must complete the display screen equipment training' may be true, but it definitely doesn't mean that all students will.

— let's see if we can make use of it to make claims about the structure of borderline cases (and non-borderline cases).

We can start to develop our system of logic for this purpose in more or less exactly the same way as we would approach the systems of modal logic just discussed. Instead of operators '$\Diamond$' and '$\Box$', we'll instead add two sentence operators 'B' and 'D' to the language of propositional logic (or to the language of first-order logic, which we'll come to in a moment), to represent the expressions 'borderline' and 'definite' respectively. This allows us to form sentences 'BA' and 'DA', representing that 'A' has borderline status or is definite.

This time 'B' and 'D' will not be duals, however. Recall that an operator 'O' is a dual of an operator 'P' when OA is equivalent to $\neg P \neg A$. If we were to treat 'borderline' and 'definite' as duals, then, we would expect that some sentence had borderline status if and only if it was not definitely not true. But this doesn't seem quite right: newborns are definitely young, and so they're not definitely *not* young, yet we definitely(!) shouldn't say that they're borderline cases of 'young'. Our notion of 'borderline' should therefore capture the idea that if an object is a borderline case of some expression, that expression neither definitely applies to it *nor definitely fails to apply*. One natural way to do this is to take 'D' as our primitive and define 'B' so that BA is equivalent to $\neg DA \wedge \neg D \neg A$.[9] A consequence of this is that, while the 'D' operator remains an analogue of '$\Box$' (the 'necessarily' operator) from before, 'B' is not an analogue of the '$\Diamond$' ('possibly') operator. Rather, it plays the same role that a 'contingency' operator plays in some systems of modal logic. To say that it's contingent whether Violet Brown will be the world's oldest woman for another two years is to say that it's neither necessary that she will nor necessary that she won't (see for example Montgomery and Routley (1966)), and we're likewise saying here that if teal is borderline blue, it's neither definite that it's blue nor definite that it isn't.

[9]This is the most common way of defining these two operators. See for example Bobzien (2013, 3).

In contrast, here's another way we can define our new operators. Instead of taking 'D' to be our primitive and defining 'B' in terms of it, we could start with 'B' and define 'D' in terms of that. Here's how we could approach this. Once we have our borderline cases fixed (where it's possible to fix them), all the cases that remain are broadly 'definite' cases. So far, this gives us that if an object isn't a borderline case of (say) 'tall', it's either definitely tall or definitely not tall. How do we distinguish these two kinds of case? The natural thing to point out that separates them is that definitely tall objects are tall, while objects which are definitely not tall are not tall. We have our definition of 'definite' in terms of 'borderline', then: an object is definitely tall when it's tall and not a borderline case of 'tall', and so DA will be equivalent to $A \wedge \neg BA$.

Whichever way we define these two operators, there are two natural claims that we should try to capture straight away, since neither are really up for debate. The first is that if some claim is definitely true, then it's true. The second is that if a sentence has borderline status, its negation also has borderline status. In defence of the first claim, I can only really say that putting 'definitely' in front of a sentence seems only to reinforce it in some way, and in doing so makes a stronger claim than the original sentence, whether we mean that it's clear from our perspective or that it's determinate in some non-epistemic sense. For the second claim to be false, we'd need there to be some possible cases in which a sentence was borderline true and yet its negation was either definitely true or definitely false. But it seems that there could be no such case: the definite truth or falsity of that sentence's negation would presumably give a definite truth-value to the sentence itself.

With this in mind, it turns out it actually doesn't matter which way of defining 'B' and 'D' we choose. I'll now sketch some (informal) arguments to show this. Note that these arguments make use of some 'classical' assumptions — in particular, disjunctive syllogism, the law of non-contradiction, and De Morgan laws.

Let's suppose we start with 'D' as a primitive and take BA to be equivalent to $\neg DA \wedge \neg D\neg A$. Then we can show that DA is equivalent to $A \wedge \neg BA$, as it is when we start with 'B' as primitive. We start by adding in the assumption that if DA is true, A is too. Then, supposing DA is true, we get that A is true, and that BA is false, since the truth of BA would require the falsity of DA. Putting this together, DA entails $A \wedge \neg BA$. Likewise, supposing $A \wedge \neg BA$ is true, we know from $\neg BA$ (and the definition of 'B') that either DA or $D\neg A$ is true, but also that $D\neg A$ can't be true since that would mean that $\neg A$ was true, contradicting A. So from $A \wedge \neg BA$ we get DA.

Now suppose we're starting with 'B' as primitive and take DA to be equivalent to $A \wedge \neg BA$. Then we can show that BA is equivalent to $\neg DA \wedge \neg D\neg A$. We start *this* by adding in the assumption that if a sentence A has borderline status, its negation does too. Now, supposing BA is true, we get that $B\neg A$ is true. From this and the definition of 'D', we can infer that both DA and $D\neg A$ must be false, since they could only be true if BA and $B\neg A$ were false, respectively, and so from BA we can infer $\neg DA \wedge \neg D\neg A$. Going in the other direction, if we suppose that $\neg DA \wedge \neg D\neg A$ is true, we can infer that both $A \wedge \neg BA$ and $\neg A \wedge \neg B\neg A$ are false using the definition of 'D'. Since BA is equivalent to $B\neg A$, this amounts to inferring that both $A \wedge \neg BA$ and $\neg A \wedge \neg BA$ are false. In turn, we can infer from this that at least one of A and $\neg BA$, and at least one of $\neg A$ and $\neg BA$ must be false. Noting that both A and $\neg A$ can't both be false (sorry dialetheists!), we can see that $\neg BA$ must be false, and so BA must be true.

Starting from either definition, then, we can arrive at the other. The difference is that if we start with 'D' as a primitive, we need to add as an extra assumption that if a sentence is definitely true, it's true, but get the fact that if a sentence has borderline status, then its negation does too, for free. And the opposite is true if we start with 'B' as a primitive. There's therefore not much to recommend one starting place over the other, then, since we're going to make sure that both of these facts come out as true anyway.[10]

3.3 Approaching Questions About 'B' and 'D'

With our language in place, let's look at how we should answer questions about the structure of different categories of borderline cases. Broadly speaking, we want to give an account of the logic which governs our new language that best captures that structure. It's perhaps best to give some idea of how we could do this by giving some examples of the kinds of claim we might make in doing so. We can consider two different frameworks for doing this — we can approach the task of giving a logic for 'B' and 'D' syntactically or semantically, in much the same way as I outlined before when looking at modal logics more broadly.

Let's start with a straightforward case — suppose we want our logic to make it so that it's always the case that if a sentence A is definitely true, then A is true. On the syntactic approach, we're trying to give a set of axioms and some rules for transforming certain kinds of sentences into others. So an appropriate way to capture this claim syntactically could be to introduce

(T') $DA \rightarrow A$

as an axiom.

What about the semantic approach? We'll need a bit more of a framework in place to reach the same goal. Just as before, we'll introduce models which interpret our language: ordered triples $< W, R, P >$, with W as a set of worlds, R as a relation that can

[10] As I pointed out above, I've only shown this using distinctively classical rules of inference. There may be similar arguments that only make use of non-classical rules, but even if a specific logic didn't contain appropriate rules to demonstrate this, we could just as well introduce either 'B' or 'D' as primitive, stipulate that the two are interdefineable in the way I outlined above, and indeed stipulate the two further axioms saying that if BA is true, then so is $B\neg A$, and that if DA is true, then so is A. None of these stipulations reflect assumptions that are really contested in the literature on vagueness — introducing any of them as stipulations just makes our logic slightly less elegant.

hold between worlds, and P as a function which gives truth-values to atomic sentences at worlds. On this approach, as with the syntactic approach, we can freely borrow resources from systems of modal logic that have already been developed. Accordingly, we can give our truth-conditions for 'D' in a directly parallel way to those for '$\Box$': DA will be true at a world α in a model M if and only if A is true in every world β in M such that $R\alpha\beta$.[11] Having established all of that, we can make it so that if a sentence is definitely true, then it's true, by stipulating that the R relation is reflexive (that is, every world can see itself).

Two interesting questions at this point: when looking at modal logic in terms of necessity and possibility, it was perhaps intuitive enough to think of the different worlds in a model as different ways things could have been, but how should we think of the worlds in our models when we're looking at definiteness and borderline status? And, for that matter, what does the R-relation represent?

Some theories of vagueness have ready answers to these questions. To give an example, supervaluationists (see for example Fine (1975), Keefe (2000)) and epistemicists (for example Williamson (1999)) typically take the worlds in their models to represent different ways of precisifying the language — some ways will make my parents old, and others won't. Supervaluationists take the R-relation to represent 'admissibility'; if a world in a model can see another, this means that it considers the world it can see to be an acceptable way of precisifying the language.[12] Put together, we get an interpretation of 'D' where a sentence is definitely true (at a world) when it's true on any acceptable way of making the language precise (according to that world).

Not all theories have explicit or obvious answers to these questions. We'll see in chapter 3 that Bobzien's approach is syntactic, for example, and so gives little focus to

[11]Given this definition and our definition of 'B', 'BA' will be true in a world α in a model M if and only if it's neither the case that A is true in every world β in M such that $R\alpha\beta$ nor that A is false in every world β such that $R\alpha\beta$.

[12]In contrast to this, epistemicists take the R-relation to represent indiscriminability.

semantic questions (which is not to say that plausible answers aren't available). I'm hoping in what follows to be relatively neutral on these questions of interpretation so that the points we can get to about higher-order vagueness are sufficiently general, and don't just apply to one or two particular theories.

3.4 Some Logical Starting Points

To finish this chapter, let's briefly make a start on developing a logic for 'B' and 'D' with some more or less uncontroversial elements we might include in our logic, before we go on to some more controversial details in the following chapter. We saw already that (T′) $DA \to A$ seems like a fairly obvious axiom we might endorse. Another natural thought that we might want to capture is that vague expressions are 'monotonic': roughly, that if a vague expression applies to an object, it applies to any object that the expression applies to 'more' or to a greater degree. So, for example, if someone is tall, anyone taller than them will be, too. Likewise if some garment can be said to be a pair of trousers, it would still be a pair of trousers if it were slightly longer.[13]

At this point it's necessary to expand our basic language. We'll add in atomic predicates ($F, G, \dots$), variables ($x_n, x_m, \dots$), quantifiers ($\forall, \exists$), names ($a_1, a_2, \dots$), and the identity sign (=). These give us a richer framework in which to look at claims about which expressions apply to which objects. For example, we can now express

[13]This last example shows that the idea of an expression applying to an object 'more' or 'less', or to a greater or lesser degree, is intended to be interpreted quite loosely, and doesn't just apply to 'graded' terms like 'tall'. To make an object 'more tall' is just to make it taller, but it's at least not grammatical to talk about a garment being 'more trousers'. Still, the idea is that if you have a garment which is a pair of trousers, you can make it 'more trousers', or trousers to a greater degree, in some informal sense by making the legs slightly longer.

Also worth noting here is that the only monotonicity principles we're going to make use of in this thesis are ones involving *atomic* vague predicates — the kind that are represented by 'F', 'G', and so on, in our language — and atomic vague predicates with a string consisting exclusively of 'D's and '$\neg$'s in front of them, such as 'DDF' or '$DDD\neg G$'. This is because being more or less 'borderline red', for example, isn't just a matter of being more red or being less red, since some shades can be made less 'borderline red' by being made more red, while others can be made less 'borderline red' by being made less red.

that some objects are tall and some aren't: taking 'F' to represent 'tall', we'll take '$\exists x_n F x_n \wedge \exists x_m \neg F x_m$' to be true. But we don't have enough resources yet to capture the 'monotonicity' claim. For that we need some notion of expressions applying 'more' or 'less', and so we'll introduce a relation '$>_F$', which will take two arguments, each of which can be a name or a variable. This will 'order' the objects we're interested in, in relation to a predicate F — for example, '$x_n >_F x_m$' will signify that x_n is 'more' F than x_m.

We can now state our monotonicity claim. We'll introduce the following as an axiom:

(M1) $\forall x_n \forall x_m ((F x_n \wedge x_n >_F x_m) \rightarrow F x_m)$.

We can read this as saying that if some object is F and some object is more F than it, then that other object is F as well. There's a negative version of this that we should state too, going in the 'other direction':

(M2) $\forall x_n \forall x_m ((\neg F x_n \wedge x_n >_F x_m) \rightarrow \neg F x_m)$;

if an object isn't F, any object less F than it isn't F either.

Having officially introduced all of the new vocabulary, I'd now like to pare things down somewhat. We'll stick to just one atomic predicate (F), and now we can just use '$>$' in place of '$>_F$'. The motivation for doing this is that what we're really interested in is the way borderline and non-borderline cases at different orders are structured with respect to some arbitrary vague expression. We'll also restrict the domain of objects that we're interested in to a finite set of objects, representing a gradual change from the expression (corresponding to 'F') applying to its not applying — a succession of tiny changes in height in the case of 'tall', or a tiny change in age for 'old', starting from uncontroversial positive cases and ending with uncontroversially negative cases. Since this set of objects is the one we're interested in, we can go further and stipulate that the members of a subset of the names in our language uniquely pick out one of each stage in such a change, and that those objects are ordered according to how much the relevant expression applies to them. So, for example, when F is interpreted as 'tall', a set of the names in our language can pick out a set of objects starting with someone who is 8ft tall and ending with someone who is 4ft tall. The 8ft tall person will be first in the ordering, and will be succeeded by successively shorter people. In this case, we'll say that the 8ft person will be picked out by 'a_1', and the 4ft person will be picked out by 'a_n', where n is the number of objects in the domain. Accordingly, it's going to be true that $\forall x_n(x_n \neq a_1 \rightarrow a_1 > x_n)$ and that $\forall x_m(x_m \neq a_n \rightarrow x_m > a_n)$: nothing is more tall than a_1, and everything is taller than a_n.

To finish describing this ordering, we're going to need a way to capture the idea of a 'successor' in the ordering. An object's successor (with respect to a predicate F) will be the object which is less F than it by the smallest amount. We'll say that x_m is x_n's successor (we'll write this as '$x_m = x_{n+1}$') when $x_n > x_m$ and $\neg\exists x_l(x_n > x_l \wedge x_l > x_m)$ are both true. On this way of spelling things out, every object (in our restricted set of objects) is going to have a unique successor, except the last in the ordering.

4 Conclusion

Where have we got to so far? We've settled on a provisional characterisation of vagueness on which all vague terms are epistemically tolerant, have (epistemic) borderline cases, and exhibit higher-order vagueness in both the sense that we broadly cannot give a theory of vagueness in precise terms, and in the sense that they have nth-order borderline cases for all $n > 0$. From there we have set ourselves a (correspondingly provisional) task — to describe the structural relations that hold between borderline and definite cases at different orders — and we now have a set of logical tools to express whatever answer we might develop.

Chapter 2

Paradoxes of Higher-order Vagueness

In the previous chapter I introduced some of the key concepts and questions that form the focus of this thesis, as well as some of the tools we need to properly investigate them. In particular, I introduced the notion of higher-order vagueness, and gave a very broad basis from which we could try to develop a logic to describe its structure. This chapter is going to examine how we can use this logic to show some problematic consequences of the assumption that there is higher-order vagueness in the first place.

The broad plan is to demonstrate a number of different ways in which the assumption that there is (radical) higher-order vagueness can be shown to lead to contradictions, looking at arguments presented by a variety of authors. Each of these arguments requires some additional assumptions — special rules of inference, or extra claims about the structure of higher-order vagueness. This chapter will show how these extra assumptions can be whittled down while still generating the desired result, ultimately arriving at my own formulation of how this can be done, which I take to require the fewest resources. I will conclude with a discussion of how much of a problem all of this is for the claim that there is radical higher-order vagueness, engaging in particular with Delia Graff Fara's suggestion that the arguments presented fail to really demonstrate that higher-order vagueness is contradictory.

1 Borderline Cases and Gap Principles

I introduced the following notion of higher-order vagueness in the previous chapter: an expression exhibits (radical) higher-order vagueness when it has nth-order borderline cases for all n. So, for example, if 'red' is (radically)[1] higher-order vague, there are borderline cases of 'red', borderline borderline cases of 'red', borderline borderline borderline cases, and so on. A natural way for the logic I introduced in the previous chapter to capture the claim that some predicate 'F' is higher-order vague is therefore to say that

(BLn) $\exists x_m(B^n F x_m)$

is true for all $n > 0$, where 'D^n' and 'B^n' stand for strings of n 'D's or 'B's respectively.

Some of the authors discussing higher-order vagueness that I'm going to engage with in this chapter look at things a little differently (with the exception of Shapiro, who we'll come to near the end), taking higher-order vagueness to involve the truth of what we can call 'gap principles' (as in Fara 2004 — so called because they express that there is a 'gap' between definite positive cases and definite negative cases of some kind) or 'no sharp boundaries' claims (as in for example Wright 1992, 2009, Zardini 2013). These have the form

(NSBn) $\neg\exists x_m(DD^n F x_m \wedge D\neg D^n F x_{m+1})$.

To give an example,

(NSB0) $\neg\exists x_m(DF x_m \wedge D\neg F x_{m+1})$,

effectively states that there are no definite (that is, non-borderline) cases of F that could be made into definite cases of *not-F* by a marginal change in its F-ness. Thinking about this in terms of knowledge, and interpreting 'DA' as 'it's possible to know that A', to give one way of understanding this, we can interpret (NSB0) as a way to capture

[1] I'm going to mostly drop the word 'radical' in the rest of this chapter for ease of reading. I'll make it clear when I mean higher-order vagueness more generally.

the claim that it's not possible to know that some bucket contains red paint and yet know that it would have non-red paint in it with a drop of white added to it.[2]

We encountered an informal version of (NSB0) earlier, when we were looking at the relationship between epistemic tolerance (that is, expressions' lacking knowable sharp boundaries) and epistemic borderline cases. To give an example at the next 'order' up, the following is supposed to capture 'second-order' vagueness:

(NSB1) $\neg\exists x_n(DDFx_n \wedge D\neg DFx_{n+1})$;

there are no definitely definite cases of the relevant F followed by definitely *not*-definitely-F cases. In other words, there's no 'sharp boundary' between the non-borderline cases and the borderline cases.[3] The authors who interpret higher-order vagueness using these gap principles take expressions to be radically higher-order vague when (NSBn) is true for all $n \geq 0$. What this amounts to is that there are no definite definite cases of the relevant expression, such as 'red', at any order, which are definitely not followed by definite cases (at the same order) of that expression.

Perhaps it looks like these authors and I are talking about two quite different things: gaps in the ordering, and borderline cases. But we can show that this difference is only superficial for our purposes, at least, because we can show that a predicate's higher-order vagueness in the sense I'm talking about entails its higher-order vagueness in the sense that *they're* talking about. There's an intuitive sense in which this connection seems to hold: if there are borderline cases of some expression, say, 'red', then there must be no boundary between definitely red objects and definitely not red ones, since the borderline cases must make up a 'buffer zone' between those two kinds of object, and likewise borderline borderline cases of 'red' would form a buffer zone between

[2]We can think about this in other terms besides knowledge, too, if we interpret 'borderline' and 'definite' in different ways. For example, if 'DA' were to represent the claim 'A is true', then (NSB0) would say that there's no bucket which can truly be said to contain red paint which could be truly said not to contain red paint if it had a drop of white added to it.

[3]This is actually a slight oversimplification. If (NSB1) is true, there's also no sharp boundary between definite cases of F and definite cases of 'not-F'.

definitely definite cases of 'red' and definite borderline cases of 'red', and so on. But more formally, what we need to show is that each instance of (BLn) (where $n > 0$) entails the corresponding (NSBn-1) — this will show that all the instances of (BLn) that constitute higher-order vagueness in my preferred sense entail all of the instances of (NSBn) that constitute higher-order vagueness in the 'gap principle' sense.

To show this, we need to help ourselves to some monotonicity claims (we saw some of these in the last chapter — recall that '$x > y$' here says that x is 'more' F than y):

(M*n) $\forall x_m \forall x_l((x_m > x_l \wedge DD^n F x_l) \rightarrow DD^n F x_m)$, and

(M*¬n) $\forall x_m \forall x_l((x_m > x_l \wedge D\neg D^n F x_m) \rightarrow D\neg D^n F x_l)$

need to be true for all $n \geq 0$. It should be uncontroversial that any instance of (M*n) is true — if an object is definitely red at any particular order, anything more red than it should be definitely red at that order too. Instances of (M*¬n) are less obviously true, but the thought is that if some object (definitely) isn't definitely red at some order, making that object *less* red isn't going to make it any closer to definitely red at that order, and so any object less red than that object will definitely not be definitely red at that order either.

With this in mind, let's now show that if (BLn) is true for every $n > 0$, then (NSBn) is true for every $n \geq 0$. We'll start by assuming, for some arbitrary $n > 0$ that (BLn) — $\exists x_m(B^n F x_m)$ — and $\exists x_m(DD^{n-1} F x_m \wedge D\neg D^{n-1} F x_{m+1})$ are true, in order to derive a contradiction, and therefore establish the second claim's negation, $\neg\exists x_m(DD^{n-1} F x_m \wedge D\neg D^{n-1} F x_{m+1})$, that is, (NSBn-1). Straight away we can infer $B^n F a_m$ from (BLn) — there's an object, a_m, which is an nth-order borderline case of F. We're ultimately going to derive a contradiction by showing that this object a_m can't be an nth-order borderline case after all. To do this, we'll first infer from $\exists x_m(DD^{n-1} F x_m \wedge D\neg D^{n-1} F x_{m+1})$ that there's some object, a_l, such that $DD^{n-1} F a_l \wedge D\neg D^{n-1} F a_{l+1}$. From there, we use our monotonicity principles from earlier to infer both $\forall x_m(x_m > a_l \rightarrow DD^{n-1} F x_m)$ and $\forall x_m(a_{l+1} > x_m \rightarrow D\neg D^{n-1} F x_m)$ — we know that a_l is definitely an $n - 1$th-order definite case of F, so everything more F than it definitely is too, and we know that a_{l+1} is definitely *not* an $n-1$th-order definite case of F, so we know that anything less F than it definitely isn't either. We also know that any object in our ordering is either more F than a_l, the same object as either a_l or a_{l+1} themselves, or less F than a_{l+1}, as a_l is more F than a_{l+1} and there are no objects between a_l and a_{l+1} in our ordering. So, given that a_l is definitely an $n-1$th-order definite case of F, and a_{l+1} is definitely not an $n-1$th-order definite case of F, the two claims we inferred from the monotonicity principles tell us that every object in the ordering — including a_m — is either definitely an $n-1$th-order definite case of F or definitely not an $n-1$th-order definite case of F.

We need to do a little more work from here to show that a_m isn't an nth-order borderline case. To do this, we need to show that being either definitely an $n-1$th-order definite case of F or definitely not an $n-1$th-order definite case of F entails not being an nth-order borderline case. We'll look at each case in turn. First, suppose a_m is definitely an $n-1$th-order definite case of F. That is, suppose $DD^{n-1}Fa_m$. Here's the strategy for showing that this entails $\neg B^n Fa_m$. $DD^{n-1}Fa_m$ entails $\neg BD^{n-1}Fa_m$ (that is, $\neg BDD^{n-2}Fa_m$) by the definition of 'D'. From this we can infer $\neg B\neg BD^{n-2}Fa_m$, using the fact that DA entails $\neg BA$. Then, noting that $B\neg A$ is equivalent to BA, we infer $\neg BBD^{n-2}Fa_m$. We can then repeat this process moving inwards — the next iteration goes from $\neg BBD^{n-2}Fa_m$ to $\neg BB\neg BD^{n-3}Fa_m$ and then to $\neg BBBD^{n-3}Fa_m$ — gradually converting 'D's to 'B's until we're left with $\neg B^n Fa_m$, showing that a_m is not an nth-order borderline case.

We can apply similar reasoning to the case where a_m is definitely not an $n-1$th-order definite case of F. When $D\neg D^{n-1}Fa_m$ is true, we can infer $\neg B\neg D^{n-1}Fa_m$ in the same way as before. Here we add in an extra step, using the equivalence of $B\neg A$ and BA to infer $\neg BD^{n-1}Fa_m$ from this, *then* $\neg B\neg BD^{n-2}Fa_m$, and finally $\neg BBD^{n-2}Fa_m$, giving us a point from which we can apply the same reasoning again. This process likewise eventually leads to the conclusion $\neg B^n Fa_m$, and so whether a_m is definitely an $n-1$th-order definite case or not, it's not an nth-order borderline case, contradicting our original assumption, and so showing that an arbitrary instance of (BLn) entails $\neg\exists x_m(DD^{n-1}Fx_m \wedge D\neg D^{n-1}Fx_{m+1})$, that is, (NSBn-1). What we've shown, then, is that if an expression exhibits higher-order *borderline* vagueness, then there are no definitely definite cases (at any order) of that expression which can be changed in a (relevant) marginal way into definitely *not* definite cases (at that order) of that expression.

2 Paradoxes of Higher-order Vagueness

2.1 Version 1: Wright

Having shown the connection between our working characterisation of higher-order vagueness in terms of borderline cases and a characterisation in terms of gap principles, let's now move on to see how we can cause trouble for the very idea that there might be higher-order vagueness (in either sense) at all, starting with an argument from Wright.

Wright begins by saying that we should endorse the following rule of inference:[4]
(DEF) If $DA_1, \ldots, DA_n$ entail C, then $DA_1, \ldots, DA_n$ entail DC.[5]
How should we understand this? Wright explicitly claims that this does not say that 'whatever is a consequence of a set of propositions each of which is definitely true is itself definitely true' (1992, 131-2) — that is, it does not say that if something is a consequence of a set of definitely true sentences, then its definitised form is also a consequence of those sentences. Instead, Wright takes it to say that 'when a proposition of the form: it is definitely the case that P, is true, it cannot be less than definitely

[4]Note, by the way, that we're approaching things *syntactically* here.
[5]I've borrowed elements from the styles of formatting this rule used by Edgington (1993, 194) and Heck (1993, 202) as I think this way of displaying rules makes things a bit clearer.

true' (1992, 131-2) — that is, if a sentence of the form 'definitely C' is true, it must also be definitely true. I have to admit here that I don't quite see this. As it appears in the proof that Wright goes on to offer, (DEF) plays exactly the former role, and not the latter, in its second instance.[6] Though it's true enough that a *consequence* of using (DEF) as a rule of inference is that if a sentence of the form 'definitely C' is true, it must also be definitely true. To see this, we just note that DA entails DA, and so applying (DEF) gives us that DA also entails DDA.

Let's look now at Wright's argument. Suppose that there is second-order vagueness — that is, suppose (NSB1) is true — and that we endorse the (DEF) rule. Wright shows a troubling consequence of these two assumptions as follows (1992, 131). Consider the definitised form of (NSB1), $D\neg\exists x_n(DDFx_n \wedge D\neg DFx_{n+1})$. We ultimately want to show that this entails the troubling claim $\forall x_n(D\neg DFx_{n+1} \rightarrow D\neg DFx_n)$: if an object is definitely not a definite case of F, its predecessor definitely isn't either. So we'll start by assuming the antecedent of this claim, $D\neg DFx_{n+1}$, that some arbitrary object's successor is definitely not a definite case of F. Such an object is either a definite case of not-F or is a borderline case of F. We want to show from this that $D\neg DFx_n$ holds, so we assume DFx_n to try to get a contradiction out of it (the negated sentence we'll infer from which we'll then definitise using (DEF)). We'll start towards this contradiction by applying the (DEF) rule to our assumption that DF_x, giving $DDFx_n$. We assumed $D\neg DFx_{n+1}$ before, so now we can introduce an 'and', giving $DDFx_n \wedge D\neg DFx_{n+1}$. But the definitised version of (NSB1) entails its own non-definitised form, $\neg\exists x_n(DDFx_n \wedge D\neg DFx_{n+1})$, which contradicts this previous claim. The assumption DFx_n therefore leads to a contradiction, and so $\neg DFx_n$ must be true. But since this follows from the two definitised assumptions $D\neg\exists x_n(DDFx_n \wedge D\neg DFx_{n+1})$ and DFx_n, we can apply the (DEF) rule again to get $D\neg DFx_n$, which is what we wanted to estab-

[6]Specifically I have in mind here the move Wright makes in using (DEF) to infer $D\neg DFx_n$ from $\neg DFx_n$ in moving from (7) to (8) in the proof detailed below.

lish from the assumption that $D\neg DFx_{n+1}$, showing that $\forall x_n(D\neg DFx_{n+1} \rightarrow D\neg DFx_n)$ must follow from $D\neg\exists x_n(DDFx_n \wedge D\neg DFx_{n+1})$.

Formally, Wright presents the argument like this (1992, 131):[7]

1	(1)	$D\neg\exists x_n(DDFx_n \wedge D\neg DFx_{n+1})$	Ass.
2	(2)	$D\neg DFx_{n+1}$	Ass. (for CP)
3	(3)	DFx_n	Ass. (for RAA)
3	(4)	$DDFx_n$	3, DEF
2,3	(5)	$\exists x_n(DDFx_n \wedge D\neg DFx_{n+1})$	2, 4, ∃-intro
1	(6)	$\neg\exists x_n(DDFx_n \wedge D\neg DFx_{n+1})$	1, (T′)
1,2	(7)	$\neg DFx_n$	3, 5, 6, RAA
1,2	(8)	$D\neg DFx_n$	7, DEF
1	(9)	$D\neg DFx_{n+1} \rightarrow D\neg DFx_n$	2, 8, CP

What's the significance of the claim that this argument establishes, that the definitised form of (NSB1) entails $\forall x_n(D\neg DFx_{n+1} \rightarrow D\neg DFx_n)$? For one, it seems that if we accept (NSB1), we should probably say that it's definitely true: whether it's acceptable or not, it doesn't seem to have borderline status. And it's worrying if this entails $\forall x_n(D\neg DFx_{n+1} \rightarrow D\neg DFx_n)$ — we know that the last object in our ordering, a_n, is as far from F (whatever F may be) as it gets — it's incredibly short when we're looking at 'tall', and really quite old when we're looking at 'young' — so $D\neg DFa_n$ looks very plausible, and yet this is going to let us infer (by successive applications of the sentence that Wright proves) $D\neg DFa_{n-1}$, $D\neg DFa_{n-2}$, and so on. All this is going to lead us to saying that a_1 itself, the first object in our ordering, is definitely not definitely F, surely flying in the face of common sense in saying, for example, that a person with no hairs on their head isn't definitely bald after all.

[7] I've made some minor cosmetic alterations to this. The final step from line (9) to $\forall x_n(D\neg DFx_{n+1} \rightarrow D\neg DFx_n)$ is also left implicit here, as it is in Wright.

This looks like a result we want to avoid. And as Wright points out, this problem is distinctively a problem for higher orders of vagueness (1992, 131-2): if we start from $D\neg\exists x_n(DFx_n \wedge D\neg Fx_{n+1})$, an expression of the claim that there are no sharp boundaries at the 'first order', we won't be able to establish $\forall x_n(D\neg Fx_{n+1} \rightarrow D\neg Fx_n)$ in the same way. To do so, we'd start from the assumption $D\neg Fx_{n+1}$, then assume Fx_n to try to get a contradiction. But we'd be unable to apply the (DEF) rule to the assumption Fx_n, because it is not definitised, so this is as far as we could go.

We therefore have our first paradox of higher-order vagueness: from the claim that vague expressions draw no sharp boundaries at the second order (that is, that there are no definitely definite cases of that expression which could be made into definite borderline cases by a small change) and an apparently plausible rule of inference governing 'D', we get a contradiction. Does this show that higher-order vagueness is impossible?

Heck gives us some reasons to think not. Consider first the way in which (DEF) is used in Wright's argument. The argument works from the starting assumption, $D\neg\exists x_n(DDFx_n \wedge D\neg DFx_{n+1})$, to the conclusion, $\forall x_n(D\neg DFx_{n+1} \rightarrow D\neg DFx_n)$, by using a conditional proof, which establishes the truth of a conditional by assuming its antecedent and showing that its consequent must also be true. Now *within* Wright's conditional proof, we have an instance of the (DEF) rule — at line (8) in the formal version. So Wright seems to take the (DEF) rule to be acceptable within conditional proofs in general.

Heck points out that this means that the following argument is therefore also acceptable by Wright's standards: Suppose DA is true. Then, by (DEF), DDA is also true. Therefore, if DA is true, DDA is true. This is a conditional proof of $DA \rightarrow DDA$ (Heck, 1993, 203). But since we also accept

(T') $DA \rightarrow A$

as an axiom, $DDA \rightarrow DA$ is a theorem of our logic — we just substitute 'A' with 'DA'.

But this means that $DA \leftrightarrow DDA$ is a theorem, and therefore 'D's are essentially redundant when we have more than one of them: DDA reduces to DA, and $\neg DDA$ reduces to $\neg DA$. Heck says that this is just not a result we're going to accept if we're taking higher-order vagueness seriously (1993, 203-4).

If it's not obvious why we ought to reject this in order to take higher-order vagueness seriously (and I don't think it is!), Edgington points out a way in which a system of logic containing $DA \leftrightarrow DDA$ as a theorem does not take the idea of higher-order vagueness seriously. Suppose we have a borderline definite case of some predicate F — that is, suppose for some x_n that $\neg DDFx_n \wedge \neg D\neg DFx_n$. Using $DA \leftrightarrow DDA$, the left hand side of this supposition, $\neg DDFx_n$, lapses into $\neg DFx_n$, and applying (DEF) to this gives $D\neg DFx_n$, contradicting the right hand side ($\neg D\neg DFx_n$). So under such a system of logic, there could be no borderline definite cases (1993, 194), which there would presumably need to be if there *were* second-order vagueness.[8]

And anyway, why did Wright choose to endorse (DEF) and not (DEF*)?

(DEF*) If $A_1, \ldots, A_n$ entail A_m, then $A_1, \ldots, A_n$ entail DA_m.

Note that this is essentially the same as (DEF), except that it doesn't require the propositions $A_1, \ldots, A_n$ to be definitised. By analogous reasoning to before, we can see that Wright would take (DEF*) to amount to saying that a sentence which is true cannot fail to be definitely true — noting that A entails A, an application of (DEF*) gives us that A entails DA. And why would it be more plausible to say that a sentence's definite truth entails that it is definitely definitely true, than it would be to say that a sentence's truth entails its definite truth? If we were thinking of rejecting (DEF) at this point, we might reject both (DEF) and (DEF*) together, then. But Heck says we don't need to reject either (1993, 204-5).

[8]We can derive the same contradiction from the existence of borderline borderline cases since, by definition, a borderline borderline case of some expression is a borderline non-definite case, and a borderline non-definite case is also a borderline definite case.

In defending (DEF) and (DEF*) as rules of inference, Heck needs to deal with the problem raised earlier — that the (DEF) rule seems just to automatically rule out higher-order vagueness — as well as an analogous problem at the first order. For if we accept (DEF*), the following reasoning looks acceptable: Suppose some sentence A is true. Then, by (DEF*), DA is true as well. So $A \rightarrow DA$. Given (T′), we can therefore establish $A \leftrightarrow DA$, which renders the 'D' operator effectively redundant, and as a consequence allows us to infer a contradiction from $\neg DFx_n \wedge \neg D\neg Fx_n$, the statement of the existence of a borderline case: $\neg DFx_n$ will entail $\neg Fx_n$ and $\neg D\neg Fx_n$ will entail $\neg\neg Fx_n$. That there could be no borderline cases *at all* is surely a result we want to avoid.

Here's Heck's solution: if we endorse both (DEF) and (DEF*) in our logic for 'D', in order to stop our logic from collapsing into one in which $A \rightarrow DA$ or $DA \rightarrow DDA$ are just theorems, while retaining the spirit of these rules, we should allow them to be used only in *main* rather than *subordinate* proofs. In effect, this means that if we have some sentence DA or A, we can only transform it into DDA or DA respectively within a proof if that DA or A doesn't depend on any suppositions that the main conclusion of that proof doesn't also rely on (1993, 204). Consider again Wright's original argument. The only supposition that the main conclusion, $\forall x_n(D\neg DFx_{n+1} \rightarrow D\neg DFx_n)$, relies on is $D\neg\exists x_n(DDFx_n \wedge D\neg DFx_{n+1})$ — the other two suppositions, lines (2) and (3) in the formal version, are both discharged before we get to the main conclusion: we only need the supposition $D\neg DFx_{n+1}$ to show that $D\neg DFx_n$ follows from it, and we only need the supposition DFx_n to show that a contradiction follows from it in order to establish $\neg DFx_n$. And consider the argument I just gave using (DEF*), establishing $A \rightarrow DA$. That argument makes use of a conditional proof, showing that DA follows from an assumed sentence A. But ultimately the conclusion, $A \rightarrow DA$, rests on *no* assumptions: the A we supposed along the way isn't needed for the truth of $A \rightarrow DA$ itself, though it appeared to allow us indirectly to prove that $A \rightarrow DA$ was true.

Heck's strategy, then, tries to defuse Wright's original argument against higher-order vagueness and against the worrying result that (DEF) and (DEF*) give way to the (apparently) implausible $DA \leftrightarrow DDA$ and $A \leftrightarrow DA$ by judging all of these to rely on inappropriate use of the (DEF) and/or (DEF*) rules. In the case of Wright's argument, (DEF) is applied in both of its instances to sentences which the main conclusion does not depend on. And likewise in the arguments presented for $DA \rightarrow DDA$ and $A \rightarrow DA$, the (DEF) and (DEF*) rules would need to be applied *within* conditional proofs.

2.2 Version 2: Fara

Perhaps Heck's reply gets to the root of why we shouldn't just be able to use (DEF) to get a contradiction straight away from the very existence of second-order vagueness (whether we get there from gap principles as in Wright's argument, or from second-order borderline cases as illustrated by Edgington). But this by no means dissolves all issues surrounding higher-order vagueness in connection with (DEF) and (DEF*). Fara points out one such issue which emerges even if we adopt the stipulation that (DEF) can be applied within main proofs.

Fara's is the first of the remaining paradoxes that we'll see in this chapter which, rather than try to show that vagueness at any order above the first order is inherently contradictory, generates contradictions from the assumption that *radical* higher-order vagueness could be exemplified across a finite set of objects. In making use of the radical nature of higher-order vagueness in this way, Fara (as well as the authors following her) is thereby able to generate a contradiction using fewer additional assumptions.

Recall that we said earlier that we could take a predicate 'F's higher-order vagueness to involve the truth of (NSBn) $\neg\exists x_m(DD^nFx_m \wedge D\neg D^nFx_{m+1})$ for any n greater than or equal to 0. And recall too that the number of objects we're dealing with, each representing a stage in a gradual transition from, for example, tall to not-tall or red to not-red, is finite: the transition can be captured in a finite number of steps without introducing any apparent sharp boundaries. Fara's argument proceeds as follows (2004, 201):[9] a_n, the least F object that we're considering, isn't F. So $\neg Fa_n$ is true. Applying the (DEF*) rule to this, we get $D\neg Fa_n$. But

(NSB0) $\neg\exists x_n(DFx_n \wedge D\neg Fx_{n+1})$,

along with $D\neg Fa_n$, entail $\neg DFa_{n-1}$, since $D\neg Fa_n$ and DFa_{n-1} couldn't both be true without contradicting (NSB0). But then we can apply (DEF*) to $\neg DFa_{n-1}$, giving $D\neg DFa_{n-1}$. From there we can infer from

(NSB1) $\neg\exists x_n(DDFx_n \wedge D\neg DFx_{n+1})$

to $\neg DDFa_{n-2}$ in just the same way as before. We can keep going like this, applying the (DEF*) rule and then the next instance of (NSBn) 'up'. Each such iteration takes us one object closer to the most F object in our set of objects, and one order of vagueness higher. The crucial point here is that we'll run out of objects before we run out of orders of vagueness, and so eventually we'll reach, for some n, $\neg D^nFa_1$ — saying that someone with no hairs on their head is not definitely bald at *some* level of definiteness, or that there's some level of doubt about whether Dulux's Flamingo Fun is pink.

Fara then points out that, starting from the assumption Fa_1, we can apply (DEF*) as many times as necessary to reach the relevant D^nFa_1, contradicting the previously established $\neg D^nFa_1$ (2004, 201).

What is this argument supposed to show? Fara intends it to be a proof that, given a finite set of objects ordered according to how 'F' they are, at least one instance of (NSBn) is not true of that set of objects.[10] Now this is perhaps not as serious a worry for the prospects of there being higher-order vagueness as Wright's argument, even if the argument itself is actually better for relying on less controversial assumptions and

[9] Note that, although I'm using (DEF*), following Fara, we could just as well use (DEF) if we ensured that all the relevant instances of (NSBn) were prefixed by $n+1$ 'D's. As in Fara's argument, we'd also need DFa_1 and $D\neg Fa_n$ to be true.

[10] If we accept that reductio can establish falsehood and not just non-truth, it also shows that at least one such instance is *false*!

methods of proof: Fara points out that we could keep adding objects to our ordered set of objects to defer any particular instance of the paradox she presents. While we can make use of 100 instances of (NSBn) to get to a contradiction when the set we're interested in contains 100 objects, we could avoid *that particular* contradiction (and so allow the relevant instance of (NSBn) to be true) by making the set just a little more fine-grained by adding more objects. Rather than displaying a transition from 'tall' to 'not tall' inch by inch, we could go centimetre by centimetre, or millimetre by millimetre, or....

We'll come back to the significance of the fact that the argument intends to show something weaker. For now we should check whether Fara's argument shows what it sets out to: to show, from the assumptions Fa_1, $\neg Fa_n$, and $m{-}1$ instances of (NSBn), for some m, that (NSBm) is not true, only applying the (DEF*) rule to sentences which the main conclusion depends on. Certainly that is what it shows if we think that reductio and (DEF*) are acceptable rules of inference and (DEF*) has been used appropriately. And it seems that it has: (DEF*) is applied to sentences derived from $\neg Fa_n$ and successive instances of (NSBn), but crucially *not* the instance of (NSBn) that we're trying to show is not true. For at the end of our chain of reasoning using instances of (NSBn), we move from $D\neg D^m Fa_2$ to $\neg DD^m Fa_1$ using (NSBm) — the instance of (NSBn) that we're trying to prove not to be true — but we don't apply (DEF*) to $\neg DD^m Fa_1$ itself; if we had, we'd have been using (DEF*) inappropriately.

To sum up, then: Fara shows how we can generate a paradox from the claim that vague expressions exhibit higher-order vagueness across a set of incredibly small changes, when that higher-order vagueness is expressed through instances of (NSBn), and assuming either (DEF*) or (DEF) is an acceptable rule of inference.

2.3 Version 3: Zardini

A fairly blunt response to Fara's argument would be to reject the (DEF*) rule as a valid rule of inference in the first place.[11] If we do this, though, which rules of inference *should* we consider acceptable? As it stands, without (DEF*), we have some fairly obvious but not hugely illuminating claims built into our logic: axiom (T′) encodes the idea that definitely true sentence are also plain true, and we have some monotonicity claims that we can make use of. We're considering adding instances of claims such as (NSBn) and (BLn) to our theory of vagueness (though note that the suggestion is not to add either kind of claim as axioms of our system of logic itself — (NSBn) and (BLn) aren't going to apply to non-vague terms, for example), which are certainly interesting, but they don't tell us anything general about how we can manipulate our vagueness-specific operators 'B' and 'D'.

Zardini suggests a rule that we ought to introduce into our logic which is not as strong as (DEF*): the 'closure of 'D' under logical consequence'. This rule says that if a sentence C follows from a set of sentences $A_1, \ldots, A_n$, then DC follows from the definitised forms of those sentences (that is, from $DA_1, \ldots, DA_n$). Zardini doesn't actually present much reason to accept this rule, other than that adding it to our logic for 'D' gives us an analogue of the system of modal logic KT (2013, 29). Nonetheless, the rule looks plausible on the face of it. At this point we can note that it's also weaker than either (DEF) or (DEF*).[12]

[11] I mentioned in presenting Fara's argument that it could also be made to work in a similar way using the weaker (DEF) rule. We might need to reject that as well, though it's worth noting that the stipulation that comes with that form of argument, that all instances of (NSBn) that get used need to be definitised for the argument to remain valid, is significant — it in itself is going to be a point of contention in the next chapter.

[12] We can sketch proofs showing that if (DEF) or (DEF*) are in force, then 'D' is closed under entailment as follows. That is, we'll try to show that if $A_1, \ldots A_n$ entail C then $DA_1, \ldots DA_n$ entail DC using the relevant rule. So suppose $A_1, \ldots A_n$ entail C. (T′) tells us that DA_1 entails A_1, DA_2 entails A_2,..., DA_n entails A_n, and so $DA_1, \ldots DA_n$ entail $A_1, \ldots, A_n$. And since $A_1, \ldots A_n$ entail C, we can infer in turn that $DA_1, \ldots DA_n$ entail C. From there we apply either the (DEF) rule or the (DEF*) rule, giving that $DA_1, \ldots DA_n$ entails DC.

Let's suppose that 'D' is closed under consequence. Zardini offers an argument which derives a contradiction from the definitised forms of a finite number of instances of (NSBn), and the claims that the objects a_1 and a_n are definitely F at all orders of definiteness and definitely not F at all orders, respectively.

We'll start by getting a grip on the argument informally and then I'll display the formal proof. We first show that a_n's being definitely not definitely F and

(NSB1) $\neg\exists x_n(DDFx_n \wedge D\neg DFx_{n+1})$

entail that a_{n-1} is not definitely definitely F. The fact that this entailment holds is going to be subsequently exploited using the closure of 'D' under consequence. To show that this entailment holds, we assume that a_n is definitely not definitely F and that (NSB1) is true, and then assume for reductio that a_{n-1} is definitely definitely F. A contradiction follows quickly from these assumptions, since a_{n-1} being definitely definitely F and a_n being definitely not definitely F is just a counterexample to (NSB1), which we assumed was true, showing that a_{n-1} must not be definitely definitely F after all.

Now having shown that a_n's being definitely not definitely F, along with (NSB1), entail that a_{n-1} is not definitely definitely F, we can use the closure of 'D' under consequence to arrive at a more significant result. We note first that a_n is definitely not F, definitely definitely not F, and so on, and that this entails $D^{n-1}D\neg DFa_n$ — that it's definite at the $n-1$th order that a_n definitely isn't definitely F — since being definitely not F implies being not definitely F. We also have among our starting assumptions (NSB1), preceded by $n-1$ 'definitely's — as with Wright's proof, we're working with the assumption that instances of (NSBn) are not just true but definitely true, and further still that they're definitely true at all orders. From these two assumptions, the result established in the last paragraph, and the closure of 'D' under consequence, we can infer $D^{n-1}\neg DDFa_{n-1}$.

Here's a formal statement of the argument so far, adapted slightly from the way it's presented by Zardini (2013, 39):

(1)	$D^{n-1}D\neg DFa_n$	Ass.
(2)	$D^{n-1}\neg\exists x_m(DDFx_m \wedge D\neg DFx_{m+1})$	Ass.
(3)	$D\neg DFa_n$	Ass.
(4)	$\neg\exists x_m(DDFx_m \wedge D\neg DFx_{m+1})$	Ass.
(5)	$DDFa_{n-1}$	Ass. (for ¬-intro)
(6)	$DDFa_{n-1} \wedge D\neg DFa_n$	5, 3, ∧-intro
(7)	$\exists x_m(DDFx_m \wedge D\neg DFx_{m+1})$	6, ∃-intro
(8)	$\perp$	7, 4, ¬-elim
(9)	$\neg DDFa_{n-1}$	5, 8, ¬-intro
(10)	$D^{n-1}\neg DDFa_{n-1}$	1, 2, 3, 4, 9, Closure

This all makes up the building blocks of the final result. We run a succession of analogous arguments from there — here's the next one. So far we've established $D^{n-1}\neg DDFa_{n-1}$. This is equivalent to $D^{n-2}D\neg DDFa_{n-1}$. But the $n-2$-times-definitised form of

(NSB2) $\neg\exists x_m(DDDFx_m \wedge D\neg DDFx_m)$

is going to let us infer $D^{n-2}\neg DDFa_{n-2}$ from this in much the same way as before. For note that $D\neg DDFa_{n-1}$ and (NSB2) are incompatible with $DDDFa_{n-2}$ — again, $DDDFa_{n-2} \wedge D\neg DDFa_{n-1}$ would just contradict (NSB2) — so from $D\neg DDFa_{n-1}$ and (NSB2) we can infer $\neg DDDFa_{n-2}$. Now we can make use of the closure of 'D' again, this time applying it to $D^{n-2}D\neg DDFa_{n-1}$ and (NSB2), definitised to the $n-2$th order, that is, $D^{n-2}\neg\exists x_m(DDDFx_m \wedge D\neg DDFx_{m+1})$. The closure of '$D$' and the fact that $D\neg DDFa_{n-1}$ and (NSB2) entail $\neg DDDFa_{n-2}$ allow us to infer from the two claims just mentioned (the ones definitised to the $n-2$th-order) to $D^{n-2}\neg DDDFa_{n-2}$, which is equivalent to $D^{n-3}D\neg DDDFa_{n-2}$.

The formal version of this part of the argument is formalised in an analogous way to the previous part (2013, 39-40):

(1)	$D^{n-2}D\neg DFa_{n-1}$	Ass.
(2)	$D^{n-2}\neg\exists x_m(DDDFx_m \wedge D\neg DDFx_{m+1})$	Ass.
(3)	$D\neg DDFa_{n-1}$	Ass.
(4)	$\neg\exists x_m(DDDFx_m \wedge D\neg DDFx_{m+1})$	Ass.
(5)	$DDDFa_{n-2}$	Ass. (for $\neg$-intro)
(6)	$DDDFa_{n-2} \wedge D\neg DDFa_{n-1}$	5, 3, $\wedge$-intro
(7)	$\exists x_m(DDDFx_m \wedge D\neg DDFx_{m+1})$	6, $\exists$-intro
(8)	$\perp$	7, 4, $\neg$-elim
(9)	$\neg DDDFa_{n-2}$	5, 8, $\neg$-intro
(10)	$D^{n-2}\neg DDDFa_{n-2}$	1, 2, 3, 4, 9, Closure

The argument keeps going like this, each time appealing to a successive (definitised) instance of (NSBn) and the result previously established (in the next case it will be $D^{n-3}D\neg DDDFa_{n-2}$), until eventually we get to $\neg D^{n+1}Fa_1$, the claim that at some (albeit very high) order of definiteness, the first object in our ordering is not definitely F, contradicting the initial assumption we made that the most F object in our ordering (say, a person with no hairs on their head if we're thinking about 'bald') is definitely F at all orders. Stating that there is higher-order vagueness (endorsing the definitised forms of a finite number of instances of (NSBn)) therefore generates contradictions alongside some plausible background assumptions, and so we have another paradox of higher-order vagueness which requires even fewer logical resources or additional assumptions than the arguments we've considered so far.

2.4 Version 4: Shapiro

Zardini's argument uses a fairly minimal set of resources to generate the contradiction that it aims at. Nevertheless, it still requires us to adopt the closure of 'D' under logical consequence, which, if nothing else, we may just feel we've been given insufficient reason to accept in the first place: perhaps Zardini's argument just shows that it's not an acceptable rule of inference after all. What's more, for Zardini's argument to work, we need to endorse definitised forms of all of the gap principles we make use of, and again it might seem less obvious that these definitised forms are acceptable, even if we think the gap principles themselves are. But we can do without them, and Shapiro points the way, at least, to seeing how.

Shapiro's argument tries to generate a contradiction from the existence of radical higher-order vagueness (exemplified over a finite set of objects), but in contrast to the arguments we've looked at so far, he generates this contradiction using the 'borderline cases' sense of 'higher-order vagueness', rather than the other authors' 'gap principles' sense. So his argument requires it to be the case that every borderline 'category' — first-order borderline case, second-order borderline case, and so on — has at least one object in it. He also requires an additional assumption, that borderline categories do not overlap — if an object is in one category, it's not in any other.

We can reconstruct the argument like this (2005, 149-50). We know that there's at least one borderline case of 'red' at each order. Let's call the second-order borderline case of 'red' we know to exist 'a_m', and the third-order borderline case we know to exist 'a_l'. We want to show not just that a_l and a_m are distinct, but that a_l is more F than a_m. This second part is going to be needed when it comes to iterating this argument to more cases. And we can show this as follows. Third-order borderline cases lie 'on the borderline' (as Shapiro puts it) between the second-order definite cases of 'red' and the second-order borderline cases, and so a_l is neither definitely a second-order definite case nor a second-order borderline case.[13] And yet, since borderline

[13] Given this description of what a third-order borderline case is from Shapiro, it seems that what he actually means by 'third-order borderline case' is 'borderline second-order *definite* cases' rather than 'borderline second-order borderline cases'. The former category can be narrower than the latter: a borderline case of 'second-order definite case of '*not*-red'' is a borderline second-order borderline case,

categories don't overlap, and a_m is a second-order borderline case, a_m must not be a third-order borderline case (that is, a borderline second-order borderline case). And being a second-order borderline case but not a borderline second-order borderline case is just to be a definite second-order borderline case, by the definition of 'D', so a_m must be a definite second-order borderline case of 'red'. a_m must therefore not be a_l, because one is a definite second-order borderline case, and one is not definitely a second-order borderline case. What's more, since third-order borderline cases lie between second-order definite cases and second-order borderline cases, and since second-order definite cases are more F than second-order borderline cases, a_l must be more F than a_m.

To generalise this, suppose some object a_k is the nth-order borderline case we know to exist (for some arbitrary n) and an object a_j is the $n+1$th-order borderline case we know to exist. As before, we therefore know that a_j, as an $n+1$th-order borderline case, is on the borderline between the nth-order definite cases and the nth-order borderline cases, and so is not definitely an nth-order borderline case. But by the same reasoning as before, a_k *is* definitely an nth-order borderline case (otherwise the categories 'nth-order borderline case' and '$n+1$th-order borderline case' would overlap), and so is not a_j, and further, a_j must be more F than a_k because it's between the nth-order borderline cases and the nth-order definite cases, where a_k isn't.

Why is this significant? What we end up with is a potentially infinite set of borderline cases at different orders, each successively more F than the last (and so each distinct from all the ones before it). But this means that there would need to be an infinite number of distinct objects in our ordering, contradicting our basic assumption that the ordering contains a finite number of objects, and so the assumption that

but need not be a borderline case of 'second-order definite case of 'red'', and need not be found between the second-order definite cases of 'red' and the second-order borderline cases of 'red'. This point applies just as much to the iterations of this term to different orders in Shapiro's argument.

there is higher-order vagueness which could be exemplified across this ordering is contradictory.

Let's not spend too long dwelling on this argument. While it doesn't require the resources of Wright, Fara, or Zardini, introducing the stipulation that borderline categories don't overlap at all is quite strange. *Definite* borderline cases of 'red' by definition couldn't be borderline borderline cases of 'red', but why should we also think that borderline cases of 'red' in general couldn't also be borderline borderline cases of 'red'? This certainly isn't going to be obvious if we're looking at *epistemic* borderline cases, for example — an object's being a borderline case of 'red' might well be consistent with the impossibility of *discovering* whether it's a borderline case. To bring to mind an example from earlier, at 6ft I'm not sure whether I'm a borderline case of 'tall', but I also think that I might be!

2.5 Version 5

I think we can do a little better than Shapiro in this last respect. While we might think that taking borderline categories not to overlap at all commits us to too much, we might find a weaker claim acceptable: that no two borderline categories overlap *completely*.[14] This is an improvement on Shapiro's stipulation because it allows for the possibility of borderline borderline cases which are also borderline cases. But it's still quite weak — crucially it doesn't allow borderline categories to be structured so that all and only the first-order borderline cases are second-order borderline cases, and third-order borderline cases, and so on. We'll see a lot more discussion about that

[14]In fact, we don't even need this strong a claim — what follows can be made to work on the assumption that for any n such that the category 'nth-order borderline case of 'red'' doesn't overlap completely with 'mth-order borderline case of 'red'' when $m < n$, there's an $l > n$ such that 'lth-order borderline case of 'red'' doesn't overlap completely with 'kth-order borderline case' when $k < l$. That is, it just needs to be guaranteed that for any borderline category that doesn't completely overlap with any of the ones 'below' it, there needs to be a borderline category 'above' it which doesn't completely overlap with any of the ones below that category.

particular structure in the next chapter, but let's assume for now it's a structure we want to rule out.

Here's the argument. We know that, for each $n > 0$, there's an nth-order borderline case in our (finite) ordering of objects for 'red'. And we know that for any n and m (where $n \neq m$), the categories 'nth-order borderline case of 'red'' and 'mth-order borderline case of 'red'' don't completely overlap. That is, if both categories have objects in them (as we know they will), there's either an nth-order borderline case of 'red' which isn't an mth-order borderline case of 'red', or there's an mth-order borderline case of 'red' which isn't an nth-order borderline case.

First, we'll name the first-order borderline case of 'red' that we know to exist 'a_m'. Now we want to expand the set of objects we know to be borderline cases of 'red' at some order, so we need some way of showing the existence of more objects which are borderline at some order. Here's a procedure for doing it. Say we've established the existence of n objects which are borderline cases at some order (so far we have 1). Now suppose that those n objects are the only ones in the whole ordering which are borderline cases at any order — that is, assume all the remaining objects are definite cases of 'red' at all orders or definite cases of 'not-red' at all orders. Now, given some arbitrary category 'mth-order borderline case of 'red'', how many ways are there to distribute mth-order borderline status and not having mth-order borderline status across the n objects we're considering? Let's say $n = 1$. Then there's 2: the one object in question can be an mth-order borderline case, or it can not be an mth-order borderline case. And if $n = 2$, there's 4: there can be two mth-order borderline cases, no mth-order borderline cases, and there are two ways in which there can be one mth-order borderline case and one object which is not an mth-order borderline case. To cut a long story short, the answer to the question of how many ways there are to distribute to objects either their having mth-order borderline status or not having it (for an arbitrary m) across n objects is 2^n. To be more exact, the number we're interested in is $2^n - 1$, because we're working under the assumption there's at least one object in each borderline category, and so the distribution of statuses which makes no object an mth-order borderline case should be ruled out as a possibility.

So far, then, we know that, relative to a set of n objects, for any m there are $2^n - 1$ ways of assigning 'mth-order borderline case of 'red'' and 'not an mth-order borderline case of 'red'' to those n objects. And we know that for each $m \leq 2^n$, there is at least one mth-order borderline case, which by supposition is exemplified by one of the n objects we know to be the only borderline cases at any order. But since there are only $2^n - 1$ ways in which each category could be distributed across the n objects, once 2^n borderline categories have been distributed over the n objects, at least two such categories must completely overlap. This contradicts the assumption that there are no such complete overlaps, and so the assumption that the n objects are the only borderline cases at any order must be false: there must be at least one more.

But we can repeat this process indefinitely for any set of n objects that we can establish the existence of, and so we can show that all of the borderline cases we know to exist cannot be exemplified across even the entire finite ordering after all.

3 How Much do the Paradoxes Show?

I'll close this chapter with some discussion of the results just established. Even if they show what they set out to, given their assumptions, do they really show that higher-order vagueness is paradoxical?

Wright's argument works quite differently from the others — it tries to show that second-order vagueness alone is contradictory, or at least that it makes the existence of a definitely not definite case of 'red' entail that there are no definite cases of 'red', and so it doesn't make use of the 'infinite' nature of higher-order vagueness. As such, the discussion I'm about to engage with doesn't really touch on Wright's argument, so I'm not going to say much more about it here. This won't affect the overall result of what I'll say in the rest of this chapter — I want to show that at least some of the paradoxes of higher-order vagueness pose genuine problems for the notion of higher-order vagueness in general, despite making use of the tension between infinite orders of vagueness and finite orderings of objects, and whether Wright's argument works does not affect this.[15]

With that to one side, let's see why the arguments I've presented may not show as much as they seem to. As I noted when presenting her argument, Fara doesn't take her argument to show that radical higher-order vagueness is inherently contradictory, only that it can't be presented as *truly* exhibited by a given finite ordering of objects (of the relevant kind). And, indeed, she points out that any particular rendition of her argument can be blocked by making that ordering more fine-grained, that is, making the small changes, according to which the objects are ordered, smaller. As a result, she says, her argument can't show that radical higher-order vagueness is contradictory relative to every possible finite ordering of objects that we could be interested in (2004, 202-3).

We can apply this sort of reasoning to the other arguments I presented, too. Zardini's argument makes use of a finite number of gap principles to show that some clearly F object isn't definitely F at some order, working 'backwards' through cases one by one, applying a pattern of reasoning showing that each such case isn't definitely F at some order. But if we add just one more object to the ordering, the argument will no longer show this — each step in the argument requires a unique gap principle to be true, and we'd 'run out' by the last step that we'd need. Shapiro's argument likewise shows how we run out of objects which must be borderline cases at some order when we only have finitely many objects to put into each category. But we could

[15]In spite of this, if Wright's argument *does* work, this only lends more support to the conclusion that higher-order vagueness is paradoxical.

also avoid any instance of this argument by increasing the number of objects in the ordering. And the last argument I presented relies on there being a number of distinct borderline cases at different orders which is higher than the number of objects in the ordering, and again, we could just increase the number of objects in the ordering until we have enough to avoid this. Any particular instance of the arguments we've considered this far could therefore be blocked just by considering orderings of objects with progressively more objects in them, and so we might argue, following Fara, that none of these arguments can show that radical higher-order vagueness is inherently contradictory, even if the assumptions they work from are correct.

Now one way in which we could respond to this would be to acknowledge that any particular argument of the sort we're looking at wouldn't work if we were considering a slightly more fine-grained ordering, but nevertheless point out that, *relative to the ordering they originally related to*, the arguments show that there is no higher-order vagueness, and that that's enough of a problem. Each of the arguments under consideration derives a contradiction from the assumption that some vague expression is vague at some particular orders, whether this is stated in terms of gap principles or claims about borderline cases. So, we could say, at least one of these orders of vagueness must not be exemplified by the ordering we're interested in. And if so, we could say that, relative to that ordering, either some sharp boundaries are in fact drawn at some order (that is, that for some n there are definite cases of 'nth-order definite case of F' followed by definite cases of 'not-nth-order definite case of F'), or there are no borderline cases of certain orders.

Fara's response to this is to say that the above presupposes that a certain kind of reductio is a valid rule of inference — we would need to say that deriving a contradiction from a gap principle (or a sentence saying that there are borderline cases at some order) is enough to show that those claims were false, rather than just *not true* (2004, 203-4). If this form of reductio isn't valid, none of the arguments we've looked at so far have shown as much as we might like about higher-order vagueness relative to any particular finite ordering of objects — we couldn't say, for example, that there *were* sharp boundaries within those orderings.

So far this isn't going to be conclusive because some theorists think that the reductio rule that Fara mentions is valid and some don't, and so quite an important claim about higher-order vagueness just seems to turn on the previous logical commitments that people have, unless the kind of result established by the paradoxes we're looking at are sufficiently serious that they make us reconsider this particular rule. Is there anything more decisive we can say?

One thing to point out in response to Fara is that there are some versions of the arguments we've looked at which can't be blocked by making the ordering more densely packed with objects because there's just no way of making things more densely packed. Perhaps we can keep making the changes to height involved with 'tall' smaller indefinitely, but consider the expression 'small natural number under 1,000'. This looks like a vague term, and we have a natural(!) set of ordered objects to consider — the natural numbers from 0 to 999. Zero is a small number if any are, and 999 shouldn't count as small since it's the biggest natural number under 1,000, and yet there seems to be no identifiable pair of numbers which marks the difference between small and non-small numbers. And since we're assuming that vagueness entails higher-order vagueness, if 'small natural number under 1,000' is vague, it's higher-order vague too. But once we've generated a paradox of higher-order vagueness for 'small natural number under 1,000', how could we make our ordering of objects more densely packed? There just are no natural numbers we could add.

Even if Fara is right that we can avoid paradoxes of higher-order vagueness in some cases by looking at objects which are changing by smaller and smaller amounts, our ability to make those changes smaller is not an essential feature of vague expressions, and so there are at least some vague expressions for which her reply will just not work. As a result, even if the reductio rule mentioned by Fara is not valid, for some vague expressions, gap principles describing vagueness at some order *could not be true* relative to those expressions, and likewise it could not be true of those expressions that they exhibited borderline cases at all orders. And if that's the case, the paradoxes of higher-order vagueness I've presented in this chapter do in fact pose a serious threat to the claim that there could be radical higher-order vagueness at all (or, at least, that all vague expressions are radically higher-order vague).

4 Conclusion

Throughout this chapter we've seen a number of ways in which the existence of radical higher-order vagueness can be shown to involve contradictions, and how this can be done using increasingly limited assumptions. The next chapter will examine the hopes of capturing the structure of higher-order vagueness while eliminating all of those assumptions, but so far I hope to have encouraged some scepticism about the possibility of doing this in general, or at least to have shown that if it is possible, it's very different to how we might have intuitively thought about it.

Chapter 3

Columnar Higher-order Vagueness

In the previous chapter we saw that there are some serious difficulties associated with capturing the structure of higher-order vagueness, understood as gap principles which apply 'at all orders', or as there being borderline cases at all orders. But I noted in explaining these difficulties that they all turn on certain further assumptions about the way higher-order vagueness is supposed to 'look'. So, could we save the idea of higher-order vagueness if we got rid of all of these assumptions? This chapter looks at how this could work, using Bobzien's 'columnar' view of higher-order vagueness as a main focus (defended in her 2010, 2011a, 2011b, 2013, 2014, 2015). Although this chapter is specifically focusing on Bobzien, and therefore can't *prove* anything about columnar views in general, the failure of this defence should give us strong reasons to reject columnar views more generally.

I'll start by outlining some essential parts of this proposal, and how it gets around the various paradoxes presented in the previous chapter. We'll then see some details of Bobzien's implementation of this view, before discussing two broad lines of criticism against it. The first is that the view relies on a badly motivated characterisation of borderline cases (itself based on a badly motivated understanding of ideal vague language use), and the second is that Bobzien's implementation of the columnar view is actually internally inconsistent.

1 Columnar Views

1.1 Key Features of Columnar Views

In the previous chapter we saw a number of ways in which we can generate paradoxes from the claim that there is higher-order vagueness, and we finished with a paradox which I took to rely on the fewest additional resources, requiring only that borderline categories don't completely overlap. Columnar views are the result of dropping that particular assumption — they take all borderline categories to overlap completely.[1] It's worth taking some time at this point to see exactly what follows from this.

One consequence of the claim that all borderline categories completely overlap is that every borderline case is also a borderline borderline case, and vice versa. So if it's not clear whether some object is red, it's also not clear whether it's clear.[2] And likewise

[1]There is some middle ground between 'no borderline categories overlap completely' and 'all borderline categories overlap completely' — namely, we could say that some, but not all, borderline categories overlap completely. But I pointed out before that the paradox I presented could be made to work using only the assumption that, for any borderline category that didn't overlap completely with every category of a lower order, there's at least one borderline category at a higher order that doesn't overlap with any borderline category at a lower order than it. Now denying *this* but saying there are *some* non-overlapping categories would commit us to saying that, beyond a certain order, any borderline category at all overlaps completely with at least one of the borderline categories 'below' it. That's a different claim from saying that all borderline categories overlap completely — there are some other, more esoteric, structures that are *consistent* with this claim. But it seems that the most sensible way to make sense of this claim is to say that, beyond some order, borderline categories all start overlapping completely with one another. And if, beyond some order, borderline categories all start overlapping completely with one another, what I have to say about this case applies just as much to the case in which *all* borderline categories overlap completely — we just think about the columns as starting 'higher up' — and so for simplicity I will stick to discussing the view on which all borderline categories completely overlap.

[2]In this chapter I'm going to use the word 'clear' essentially as a synonym for 'definite'. This is a fairly natural reading anyway, but I use it here in particular to mesh with the account that Bobzien puts forward is stated in terms of 'clarity', and she gives a special definition for this term. Bobzien also uses the letter 'C' (for 'clear') in her system of logic rather than the 'D' (for 'definite') that we've been using so far. There's nothing important about this difference so I'm going to continue to use 'D'.

it won't be clear whether it's unclear whether such a case is an unclear case of 'red', and so on. The crucial thing to note here is that, on this picture, there are no clear borderline cases of any sort because a clear borderline case is exactly a non-borderline borderline case, and this view says that there are no such things.

This may come across as counterintuitive, but perhaps it can be made to look a little more plausible at this point. I've made reference to cases previously which I took to be borderline cases — the walls of the IC for 'blue', and I thought I might be a borderline case of 'borderline tall' — but I'm not confident that everyone would agree with me about either, and so perhaps neither is a good candidate for being a clear borderline case. And yet clear non-borderline cases are easy to come by — someone with no hairs on their head is clearly not a borderline case of 'bald', and likewise one grain of sand is clearly not a borderline case of 'heap' — so if we're struggling to identify any clear borderline cases, shouldn't we be suspicious of their existence?[3] On the other hand, we might attribute these difficulties to issues surrounding context. Perhaps it's difficult to definitively identify borderline cases because this would require them to be borderline cases regardless of the context in which we're talking about them. We can lend some support to this thought by pointing out that clear cases are actually harder than it seems to identify once and for all, perhaps for this exact reason — Violet Brown is clearly old when we're talking about humans, but when we're talking about living things on Earth, she's relatively young (apparently the oldest tree alone is over 5,000 years old!). I say all of this to put the claim that there are no clear borderline cases into perspective, as it may seem a little outlandish at first; it will receive more attention soon.

The other side of the coin is that, since all borderline categories overlap completely, all non-borderline cases will be clearly non-borderline. To see why this is so, consider a clear case of 'tall' (say, someone who is 8ft tall). Such a person is not a borderline case

[3]If you think you've identified one, try asking 5 people if they agree and see what happens!

of ‘tall’. And the claim that all borderline categories completely overlap entails that if an object is not a member of one borderline category (with respect to some vague term), it is not a member of any. So this person must not be a borderline borderline case of ‘tall’. This means they’re either clearly clearly tall, or clearly clearly not tall, and so from the fact that they’re clearly tall we can infer that they’re clearly clearly tall. The same argument can be applied to show that people who are clearly *not* tall are likewise clearly clearly not tall, and so we can show that any non-borderline case will be clearly non-borderline.

Given these two features of columnar views, we can see where the term ‘columnar’ comes from. Say we have a set of shades going from Volcanic Red to Tangerine Twist in a succession of small changes. On the columnar view, we can divide them neatly into three sorts: shades which are clearly red, clearly clearly red, and so on, shades which are borderline cases of ‘red’, borderline borderline cases of ‘red’, and so on, and finally those which are clearly not red, clearly clearly not red, and so on. Between them they form three ‘columns’, thinking of the ordered shades themselves marking a ‘horizontal’ axis, and the orders of clarity (or borderlineness) marking the ‘vertical’:

$D \ldots DF$	$D \ldots DF$	$D \ldots DF$	$B \ldots BF$	$B \ldots BF$	$D \ldots D\neg F$	$D \ldots D\neg F$	$D \ldots D\neg F$
$DDDF$	$DDDF$	$DDDF$	$BBBF$	$BBBF$	$DDD\neg F$	$DDD\neg F$	$DDD\neg F$
DDF	DDF	DDF	BBF	BBF	$DD\neg F$	$DD\neg F$	$DD\neg F$
DF	DF	DF	BF	BF	$D\neg F$	$D\neg F$	$D\neg F$
1	2	3	4	5	6	7	8

Contrast this with the alternative picture, on which borderline categories at different orders can differ in which objects they contain, and likewise for categories marking different levels of clarity. This gives us what we can call, following Bobzien (2010), a ‘hierarchical’ view of higher-order vagueness — displaying this structure graphically tends to yield upside-down pyramids, with the borderline cases at the base, then second-order borderline cases (as we might expect) on the borderlines of that category, then third-order borderline cases on the borderlines of *that* category, and so on:

$DDDF$	$DDDF$	$\ldots$	$BBBF$	$DBBF$	$DDBF$	$DBB\neg F$	$BBB\neg F$	$\ldots$	$DDD\neg F$	$DDD\neg F$
DDF	DDF			BBF	DBF	$BB\neg F$			$DD\neg F$	$DD\neg F$
DF	DF				BF				$D\neg F$	$D\neg F$
1	2		3	4	5	6	7		8	9

1.2 Logic for Columns

How can we represent the structural features of columnar views using our logic of 'B' and 'D'? Bobzien has a well-developed answer to this question, so we can follow her at least in part in trying to answer it ourselves. So, following her, and approaching the question syntactically (that is, roughly, by providing axioms and rules of inference), we'll say that the two core axioms of the columnar view are:

(4) $DA \to DDA$, and

(V) $BA \to BBA$;

if something is clear, then it's clearly clear, and if something is borderline, then it's borderline borderline. We'll also treat

(T′) $DA \to A$ and

(K) $D(A \to B) \to (DA \to DB)$

as axioms, though this isn't a special feature of columnar views.

From the axioms we have so far, we can derive the converse forms of the main two, that is,

(4c) $DDA \to DA$, and

(Vc) $BBA \to BA$.

(4c) is just an instance of (T′). And suppose that $\neg(BBA \to BA)$ is true for reductio. Then $BBA \wedge \neg BA$ is true too. $\neg BA$ is equivalent to $\neg(\neg DA \wedge \neg D\neg A)$, and so $DA \vee D\neg A$ must be true. But then, applying (4) to both sides of this disjunction in turn, either

DDA or $DD\neg A$ must be true. Now, DA entails $\neg BA$, and $D\neg A$ entails $\neg B\neg A$, which in turn entails $\neg BA$. So whether DDA is true or $DD\neg A$ (and we know one must be), $D\neg BA$ must be true. This entails $\neg B\neg BA$ and so $\neg BBA$, which contradicts our earlier claim $BBA \wedge \neg BA$, and so (Vc) must be true.

Putting all of the above together, we can therefore say that

(4*) $DA \leftrightarrow DDA$, and

(V*) $BA \leftrightarrow BBA$

are theorems of the columnar view, and in turn that

(4**) $D^n A \leftrightarrow D^m A$ and

(V**) $B^n A \leftrightarrow B^m A$

are theorems for any $n > 0$ and $m > 0$.

With respect to rules of inference, the columnar view doesn't in itself seem to force many choices. In what follows, I'll make use of rules as conservatively as possible.

That's the basics of it. The full details of Bobzien's logic are interesting, but I'm not going to deal with anything much more detailed than what I've outlined so far. To give an example that will be useful shortly, however, one principle that gets validated is that if an expression has borderline cases, it's unclear whether it has borderline cases. Bobzien argues for this aspect of her logic by making use of a further assumption, that for there to *clearly be* some objects to which an expression applies, that expression must *clearly apply* to some cases.[4] Since there are no clear borderline cases under the columnar view, she infers from this that it can't be clear whether there are any borderline cases of a vague predicate, say, 'red' (if it is vague), because, on the columnar view, there are no objects to which 'borderline case of 'red'' clearly applies (Bobzien, 2015, 75-6). A consequence of the claim that it cannot ever be clear that an expression has borderline cases, given our assumption that all vague expressions

[4]This assumption seems reasonable to me, though I don't intend to present any argument for it either way: I bring it up just as a feature of Bobzien's system.

have borderline cases, is that there could be no clearly vague expressions: to find a clear vague expression would be to find an expression which clearly had borderline cases.[5]

1.3 Avoiding the Paradoxes

In introducing the columnar view, I said that one of its key attractions is that it avoids the paradoxes of higher-order vagueness, so it's worth making it explicit how exactly it does this. Bobzien has offered responses to most of the arguments presented in the previous chapter, so I'll go through replies to each of them, (critically) making use of her responses where possible.

We'll start with Wright's 1992 argument. Recall that this used the (DEF) rule (If $DA_1, \ldots, DA_n$ entail C, then $DA_1, \ldots, DA_n$ entail DC) and a definitised gap principle $D\neg\exists x_n(DDFx_n \wedge D\neg DFx_{n+1})$ to show that if an object was definitely not-definitely red, say, then an object slightly more red is also definitely not-definitely red, meaning that if there are any definitely non-definitely red objects, there are no definite cases of 'red' at all.

Bobzien's reply to this argument on behalf of the columnar view is slightly puzzling. She first claims that, besides accepting the conclusion (and so that higher-order vagueness is paradoxical), we could either reject (DEF) as a rule of inference, reject the definitised gap principle, or reject the 'second' claim that Wright uses in his argument, that there are some clear borderline cases. The way she presents this suggests that she has in mind here that rejecting this last option is to be done by rejecting that any objects are definitely not-definitely red, and this is her suggested response (2010, 22).[6]

[5]Bobzien lets the mask slip on this line sometimes — she describes 'vague' itself as 'undoubtedly' a vague predicate, for example (2010, 3), then later says she can make no judgement on it (2015, 76).

[6]Here's what she specifically says. In her formal reconstruction of Wright's argument, line (2) says '$Def(\neg Def(Fx'))$' ($D\neg DFx_{n+1}$ in our terms) and is labelled '2nd assump.'. She then goes on to

Rejecting this claim *would* make Wright's argument toothless, since even if the conclusion of the main argument, $\forall x_n(D\neg DFx_{n+1} \to D\neg DFx_n)$ were true, if there were no definitely non-definite cases of 'red' to begin with, there would be nowhere to start the chain of reasoning showing that some obviously red objects were in fact not definitely red. But I say that this is puzzling for two reasons. First, because the claim that there are no definitely non-definite cases of 'red' is *not* the same as the claim that there are no definite borderline cases of 'red'. Saying that there are no definitely non-definite cases of 'red' does *entail* that there are no definite borderline cases of 'red', but it also entails that there are no definitely definite cases of 'not red': a blue shade, say, is definitely definitely not red, and so it's surely definitely not a *definite* shade of red. So if Bobzien takes her preferred approach, she also seems forced to deny that there are definitely definite cases of 'not red', which is a strange thing to say, at least.

The second puzzling feature of Bobzien's response is that she correctly identifies a crucial distinction which would allow her to give a better response to Wright on behalf of the columnar view, and yet does not actually give that response. This crucial distinction is that Wright's argument uses a definitised, rather than non-definitised, form of a gap principle. This distinction is important because, while the columnar view tries to preserve higher-order vagueness, and so seems bound to accept that vague expressions satisfy gap principles at many orders, it can avoid saying that any of them *definitely* do.

To see how, we need to think back to the discussion of how to characterise vagueness presented in chapter one. I suggested that epistemic tolerance — roughly, that

say '[The] options we have are: either give up DEF, or give up *the definitization* of $[\neg\exists x_n(DDFx_n \wedge D\neg D\neg Fx_{n+1}]$ (the 1st assumption), or give up definite borderline cases of the first order (the 2nd assumption), or admit that higher-order vagueness is paradoxical' (2010, 22, her emphasis). The 'first assumption' she talks about here is labelled as such at line (1) of her reconstruction of Wright's argument, in the same way that the 'second assumption' is labelled. She then says that her implementation of the columnar view 'rejects the existence of clear (or definite) borderline cases. In other words, it rejects Wright's 2nd assumption. Thus, even if DEF and $[\neg\exists x_n(DDFx_n \wedge D\neg D\neg Fx_{n+1}]$ are admitted, no higher-order vagueness paradox ensues' (2010, 22).

there are no knowable sharp boundaries — is characteristic of vague expressions, and subsequently presented an argument from Greenough showing that if an expression is epistemically tolerant, it has (epistemic) borderline cases. Now, the claim that an expression is epistemically tolerant is actually just a kind of gap principle — 'D' is just interpreted epistemically — so Greenough's argument can be used in exactly the same way to show that if an expression makes a gap principle true (at some order), it also has borderline cases at the relevant order. A defender of the columnar view could use these resources to show why we should reject the definitised form of any gap principle by helping themselves to Bobzien's claim that no expressions clearly have borderline cases, and inferring from there that no expressions clearly make any gap principles true. Such an inference would be acceptable (it seems) because if an expression did clearly make a gap principle true, it would also clearly have borderline cases, and accepting the latter would just be to deny Bobzien's claim. In sum, then, columnar views can avoid Wright's conclusion by denying the only premise of his argument, the definitised form of a 'second-order' gap principle.

Moving on to Fara's formulation of the paradox, we need to take a slightly different approach because no use is made of any claims which take some expression to be *clearly* vague in any sense. Fara's argument works by applying gap principles successively to reach a contradiction from the truth of DFa_1 and $D\neg Fa_n$ — that some object is clearly red and some object much less red than it clearly isn't red, say — using the (DEF*) rule (If $A_1, \ldots, A_n$ entail C, then $A_1, \ldots, A_n$ entail DC).

Bobzien's response to Fara's version is just to reject (DEF*) as a valid rule of inference, since it is just inconsistent with the columnar view (2010, 23). If (DEF*) were a valid rule then, since $BA \vdash BA$ is presumably valid, (DEF*) would tell us that $BA \vdash DBA$ is also valid. But while the columnar view is consistent with there being borderline cases, it is not consistent with there being clear borderline cases. So

$BA \vdash DBA$ could not be valid, and so (DEF*) must in turn not be consistent with the columnar view.

Zardini's argument can be dealt with in much the same way as Wright's. His argument makes use of $D^m Fa_1$ and $D^m \neg Fa_n$ (for some arbitrarily high m), the closure of 'D' under consequence, and a number of gap principles, definitised to different degrees, to reach a contradiction. To avoid Zardini's conclusion, then, a defender of the columnar view should reject any one of these definitised gap principles, again pointing out that the truth of one of these would be inconsistent with the columnar view itself, as it would entail the existence of clear borderline cases.

Finally, both Shapiro's argument and the argument that I presented can be deflected easily by the columnar view. Shapiro's makes use of the assumption that borderline categories do not overlap at all, and the argument I presented used the assumption that these categories don't overlap completely. These assumptions are just false on the columnar view, since it takes all borderline categories to overlap completely.

1.4 No Hierarchies — Why Columns?

So far all I've shown is that the columnar view can maintain consistency in the face of various paradoxes of higher-order vagueness, while still keeping hold of the claim that vague expressions are higher-order vague. Certainly this is worth something — if there is higher-order vagueness, a theory of vagueness should presumably be able to capture it consistently. But given its apparently strange theoretical commitments, and without any substantive positive reasons to think that a *columnar* view of higher-order vagueness is true in the first place, doesn't such a view just represent an ad hoc way to cling on to the phenomenon?

I think the fairest way to answer this last question is to look at a particular defender of the columnar view — indeed, the only person to attempt a defence — and see why they think the view is true. This should give us an idea of the kind of reasoning that might lead us to accept the view in the first place, and from there we can see whether the reasoning is ad hoc after all, and indeed whether it is good reasoning by its own lights.

2 An Outline of Bobzien's Columnar View

2.1 Borderline Cases and Ideal Speakers

The key difference between hierarchical and columnar views of higher-order vagueness is in what they have to say about the relationship between different borderline categories. Obviously this may have an impact on many other things — whether we can say that there are any vague expressions at all, and what relationships hold between different kinds of clarity, for example — but what the different views have to say about borderline cases in particular ought to be the main focus of examining them. The columnar view ultimately requires us to make two bold claims about borderline cases: that there are no non-borderline borderline cases, and that there are no borderline non-borderline cases, of any expression. That's how we get all of the borderline categories to line up. So, in looking at Bobzien's view, what we're looking for is a justification of these two claims.

Now, I noted back in chapter 1 that Bobzien is eager to point out that the expression 'borderline case' is ambiguous. We can have 'not-quite' borderline cases and 'only just' borderline cases, for example — it's possible to say about these that the relevant expression does not, or does, apply to them respectively, but they are on the limits of cases that can be classified in this way. All sides of the debate can presumably agree that neither of these are especially relevant to thinking about vagueness, since under either of these definitions a borderline case has some determinate status. But a potentially more interesting notion of a 'borderline case' described by Bobzien is the notion of an 'in-between' borderline case: an object is an *in-between* borderline case when it falls between two exclusive classifications (2013, 5-6; 2015, 78). So, for example, we might think that hills are in-between borderline cases with respect to outcrops and mountains — for something to be a hill, it must first be in-between those two classifications.

Bobzien rejects the 'in-between' sense of 'borderline case' as being the sense that's relevant to vagueness. Instead, she favours an epistemic reading of the term, introducing the concept of an 'undecidable' borderline case, an object such that it's not possible to tell whether the relevant expression applies to that object or not. This is at least what Bobzien says she means by an 'undecidable' borderline case (2015, 78). We'll see later that she's a little ambivalent about what counts.

Given what we've said so far about characterising vagueness in terms of epistemic borderline cases in chapter 1, the suggestion that the kind of borderline case that's relevant to vagueness is epistemic rather than 'in-between' shouldn't come across as objectionable, as it allows us to maintain neutrality towards different theories of vagueness. But Bobzien's characterisation, in terms of 'ability to tell' has a very specific meaning in her terminology.[7]

One thought that we've encountered already is that an object's status as a borderline case (when we're thinking about vagueness, at least) depends on its resistance to investigation — just counting all the grains in some collection of sand isn't always going to tell us whether it's a heap, for example. So we might in turn think that an object's borderline status should not depend on particular speakers having carried out insufficient or poor investigation into them — Upper Hannover Street in Sheffield seems to be a borderline case of 'steep' because measuring its incline or walking up and down

[7] It's especially worth noting here that she doesn't intend her definition of 'can tell' to be a reflection of the conventional use of that expression (2010, 66).

it would be inconclusive to tell either way; if it was just that no one had bothered to do it, that wouldn't be enough. So, Bobzien says, what counts as a borderline case must be responsive only to the judgements that speakers who had properly investigated a given object would make about it (2010, 7). Any other speakers' judgements on that object could be clouded by some lack of information or thought on their part.

What counts as being able to investigate properly whether an expression applies to some object or not? Bobzien singles out three features that a speaker of (in our case) English ought to have: competence with the relevant expression, the ability to reason well (infallibly, in fact), and possession of all the information that's relevant to evaluating whether the expression applies or not (2010, 7-8). These three seem intuitive enough: a street's being borderline steep shouldn't depend on the judgement of someone who just doesn't understand the term 'steep'; someone with faulty reasoning could infer that it was (or wasn't) steep for all the wrong reasons; and someone who has never heard about it, or who has never looked especially hard at it, presumably won't make judgements about it that most accurately reflect its status.

With all this in mind, Bobzien says that the speakers whose judgements we should be interested in when asking whether some object is a borderline case of some expression are speakers who are ideal in all of the above respects: they are maximally competent (with respect to the relevant expression), rational, and informed (2010, 7-8).[8] Are there any such speakers? In some cases there presumably are, at least: you and I are probably as well-equipped as anyone else is to say that Blake Street is steep (once you've seen it, at least, or once I tell you it's the third-steepest residential street in the UK). But in others, Bobzien says, it doesn't really matter whether there are any such speakers — what matters are the judgements that those speakers *would* make if they were to exist and to be put in the position of evaluating the relevant object (2010,

[8]Bobzien calls these speakers 'CRISPs' — competent, rational, informed speakers — but this just sounds a little odd to me, so I refer to them here as 'ideal speakers'.

25). Mature redwoods are clearly tall, even when no one's looking at them, because anyone who did see one couldn't avoid that judgement when pressed.

We can now be a bit more explicit about Bobzien's interpretation of 'being able to tell'. It's possible to tell (in her sense) that some expression applies to some object when, were any speaker who was perfectly competent with that expression, perfectly rational, and maximally informed, presented with that object and asked whether that expression applied to it or not, would assert that it did, given as much time as they wanted to reach that conclusion (2010, 7). And, as far as Bobzien is concerned, an object is a borderline case of some expression just when it's not possible to tell, in the above sense, whether or not that expression applies to that object.

2.2 Ideal Speakers and the Columnar View

All of this might sound reasonable enough so far. But the difficulty comes from showing that speakers who are best equipped to make judgements about cases would in fact make judgements that reflect the columnar view. Merely defining what a borderline case is in terms of ideal speakers' assertions definitely doesn't give us enough reason to endorse the view. Bobzien therefore offers some further claims about these speakers in order to bridge this gap.

According to her, an ideal process of evaluation involves first making an initial judgement on the relevant case — perhaps you need to glance down Upper Hannover Street to make such a judgement on whether it's steep, for example — and then investigating the case further, perhaps indefinitely, perhaps making measurements or looking from different angles in the case of 'steep'. Then, if an ideal speaker is confident that the relevant expression applies to the relevant object, they will assert that it does (assuming they have been asked), and would likewise assert that it did not apply if they were confident of that instead. They might also reach no conclusion at all (2010, 7-8,11).

If these speakers do reach an answer, Bobzien says, that answer will be correct (2010, 7). The thought behind this might just be that if someone is really competent and has investigated a case properly, they shouldn't make any 'mistakes' in their assertions, since they couldn't investigate any better, and will only make assertions when they're sufficiently confident in them.

An immediate consequence of this is that ideal speakers cannot make assertions about particular cases that contradict one another. If they did, an expression could both apply and not apply to an object, which seems to be a contradiction we should avoid. Bobzien takes this a step further, saying that we should also rule out the possibility of 'communication failure' between ideal speakers. She doesn't actually say what this means, but from context she seems to mean that she wants to rule out the possibility of one ideal speaker asserting that, say, Crookes Road is steep and another speaker *failing* to assert it (when asked), even given enough time to investigate (2010, 9). Her reasons for ruling out this 'communication failure' aren't clear (leaving aside accusations of ad hoc theorising), but perhaps we could defend this claim by saying that if *some* ideal investigators can detect that Crookes Road is steep, then any of them should be able to, since they all have the same resources available to them, and are all investigating in the same ideal conditions.

To ensure both that ideal speakers don't make conflicting assertions and that they all arrive at the same conclusions, Bobzien says that they leave 'margins for error' in their assertions, drawing on similar treatments in Glanzberg (2004), and presumably from Williamson (1994) at least implicitly. The general idea here is that speakers leaving margins for error will only ever make assertions when they have 'left room' in their judgements for small discrepancies of some kind which are relevant to their judgements, but which aren't directly taken into account by the way in which they go about making those judgements. In some contexts (for example in Williamson (1994)), leaving a margin for error will mean not judging an oak tree to be exactly 18m 1cm tall (just by looking) when we know that it's around 18m tall but that we can't tell the difference between something being 18m tall and its being a centimetre taller than that — that is, when we know that there are certain discrepancies in height that 'just looking' won't be sensitive to. In Bobzien's sense, leaving a margin for error in making judgements means leaving space for discrepancies in other ideal speakers' initial reactions to particular cases.

We can be very specific about the margins for error that ideal speakers leave on Bobzien's view. What accounting for discrepancies in initial reactions amounts to is that an ideal speaker will only assert that an expression applies to an object if it is sufficiently 'obvious' in some sense — if *any* ideal speaker would have the initial reaction that it does apply.[9] We can spell this out precisely, as follows. Take one of our orderings of objects from before — say, patches going from red to not red — and take all of the relevant speakers' initial reactions to all of them. Take the largest section of the ordering in which there is variation in initial reaction, where some speakers (initially) think some objects are red and others (initially) think they aren't. Call the number of objects in this section 'e'. Under Bobzien's margin for error constraint, speakers will only *assert* that some object in the ordering is red if they'd also have the initial reaction that the shades e shades away from it on either side were red. That is, they will only assert that an object is red if they would still have the initial reaction that it was red even if it were slightly more or slightly less red (2010, 9).

Leaving this room ensures that no ideal speakers assert that any objects are red while others didn't even initially think that they were red. Assuming these speakers only get more likely to initially think that an object is red as it becomes more red,

[9]Bobzien takes ideal speakers to 'automatically' do this, but suggests that they could also achieve this by conferring with one another. (2010, 9).

disagreement in assertion is impossible once a speaker has left this room; if there is disagreement within all of those e shades on one side of an object or the other, there must not be any disagreement about the object itself, as the largest section of the ordering in which there was disagreement would then contain at least $e + 1$ objects.

A diagram might be useful here. Let's think back to our ordering containing 8 shades from earlier, going from red to not red. We can represent a possible set of reactions to these shades from 3 speakers as follows:

	1	2	3	4	5	6	7	8
S1	red	red	red	red	red	not red	not red	not red
S2	red	red	red	red	not red	not red	not red	not red
S3	red	red	red	not red	not red	not red	not red	not red

The dotted lines around shades 4 and 5 close off the largest section of this ordering in which there is disagreement, and so the relevant e for this ordering is 2. Speakers obeying a margin for error constraint can therefore only assert that a shade n is red if they would have the initial reactions that shades $n-2, n-1, n, n+1, n+2$ (where those shades exist) were red. So in this case, for example, speaker 1 could assert that shade 3 was red, but not 4 — leaving a gap of 2 objects ensures that they don't stray into the section where there is disagreement.

Bobzien says that these margin for error constraints on ideal speakers entail that they are all unanimous in their assertions — that is, if one ideal speaker would make an assertion about some case, any ideal speaker would (2010, 10). Unhelpfully, she doesn't say why she thinks this is so. In fact, this claim appears to be false. Looking at the possible set of reactions to shades that I gave above, under Bobzien's margin for error constraints it seems that, while speaker 1 could assert that shade 3 was red, speakers 2 and 3 could not, since each had the initial reaction that shade 5 was not red. In which case, it's perfectly consistent with margins for error that some ideal speakers might make assertions that not all ideal speakers would make. Note, though, that this still doesn't mean that ideal speakers can deliver *conflicting* assertions; margins for error do rule that out.

Bobzien does need it to be the case that ideal speakers are unanimous in their assertions in order to argue for axiom (4) of the columnar view, as we'll see in a moment, so is it a serious problem if margins for error cannot guarantee unanimity? Perhaps not; instead, she might modify her characterisation of ideal speakers to guarantee unanimity in a different way. For example, she entertains the idea of saying that ideal speakers could, before making assertions, communicate with one another about their initial reactions to cases (2010, 9). In this case, she could also say that they will only assert that an expression applies to an object when everyone else had the same initial reaction. This would allow speakers 1, 2, and 3 in the situation represented above to assert that patch 3 is red, for example. But even this still doesn't quite guarantee unanimity. Presumably the process by which ideal speakers decide whether an object is red isn't just a process of aggregating initial reactions, and so there's still room for any of the speakers to hesitate or change their minds. Still, let's keep going with the assumption that ideal speakers are unanimous.

How does Bobzien's more detailed characterisation of ideal speakers get us to the columnar view? Let's start with axiom (4), $DA \rightarrow DDA$. Remember that, on her approach, 'DA' means that it's possible to tell that A, and so 'DDA' means it's possible to tell that it's possible to tell that A. So to show (4) to be correct we need to show, from the supposition that any ideal speaker would assert A, that any ideal speaker would assert that any ideal speaker would assert A. Bobzien does this by first saying that if one ideal speaker would assert some sentence A, then any ideal speaker would — this is the unanimity assumption. She then says that anyone who was an ideal speaker with respect to the expression 'any ideal speaker would assert A' (that is, 'DA') could recognise some particular ideal speaker and could in turn recognise that they would assert A, and could infer from this that any ideal speaker would assert that A. They would therefore themselves assert that any ideal speaker would assert A, and since this is an arbitrary ideal speaker, we can therefore infer that any ideal speaker would assert that any ideal speaker would assert A, and so (4) is true for any A (2010, 16-7).

The route to (V), $BA \rightarrow BBA$ is a little less straightforward than this, and Bobzien ultimately tries to establish this principle in a few different ways. In her (2010) (where the rest of these details originate) she says that an ideal speaker who is confronted with a case where they are unable to rule out that the expression applies to it or does not, and so a case about which they cannot make any assertions, doesn't *actively* withhold judgement about that case. For them, she says, 'it is as if [they] have never been asked' (2010, 11). So, while these speakers cannot tell whether the relevant expression applies, they will not assert that they cannot tell. And likewise they cannot identify borderline cases by pointing to times when they withheld judgement, because '[their] consideration of the question [of whether the expression applies to the object or not] took time, and they cannot rule out that during that time, unnoticed by [them], a slight error occurred in [their] evaluation' (2010, 12) which is just big enough to prevent them from making an assertion either way. If this is so, there is no way for an ideal speaker to identify any borderline cases of any kind, and so an object's borderline status must always entail borderline borderline status. We therefore have the two key axioms of the columnar view.

3 Against Bobzien's Account of Borderline Cases

3.1 Problems with Ideal Speakers

If we're treating 'borderline cases' as picking out epistemic borderline cases, we might think that using ideal speakers as a measure of what should and should not count as a borderline case is a good idea. But we might nevertheless disagree with the way in which Bobzien characterises these speakers, leaving aside the broad strokes of ideal competence, rationality and information. Let's reflect on some of these before moving on to some more substantive objections to the view.

The first worry we might have is with the margins for error that ideal speakers are supposed to leave under Bobzien's view. Recall that these speakers will only assert that, say, some person is tall if they would have had the initial reaction that a person who was shorter than them (by a height corresponding to e) was tall, where e is the number of objects in the range of heights about which ideal speakers would differ in initial reaction about whether they were tall. While we might think that *some* margin for error might be left by ideal speakers, it is not clear why this one in particular would. For one, ideal speakers on Bobzien's view need to achieve the psychologically implausible in automatically leaving margins for error if those margins correspond exactly to the relevant e across an ordering.[10] Ordinary rational and competent speakers in the best of conditions would be doing something miraculous if they left exactly the correct margins. Now Bobzien might brush this worry to one side by saying that *ideally* all speakers would leave such margins despite limitations on *actual* linguistic practice (although she is explicit in saying that she does not think that ideal speakers are superhuman (2010, 7)), or she could appeal again to the idea of 'surveying' other ideal speakers to determine their reactions in order to ensure that any assertions came with no 'risk'. But either of these options just raise another worry. There's no reason to think that ideal speakers would only make judgements once they had made sure everyone agreed with them initially: learning that just one speaker had a different *initial reaction* to a case than the considered judgement you'd like to make would probably not stop you, or any reasonable speaker, from making that judgement

[10] To be precise, the height difference that speakers would have to be sensitive to would be the number e multiplied by the height difference between adjacent objects in the relevant set.

explicit. At least, it seems it wouldn't be unreasonable to make that judgement in spite of that disagreement.

A related worry we might have is with Bobzien's claim that (as a result of their leaving margins for error or consulting other ideal speakers) ideal speakers do not make conflicting *assertions*. This just doesn't seem to fit with actual linguistic practice. Consider an alternative picture from Wright. He says instead that, when considering cases going from, for example, tall to not tall, competent speakers will begin making more 'uncomfortable' verdicts and then become more comfortable again. He says that such speakers nevertheless *make* less comfortable assertions, if hesitantly, and that disagreement between speakers on such assertions is to be expected, and indeed that this is no mark of incompetent use. On this view, if two speakers disagree in their verdicts, or if only one reaches a verdict, this by itself is no reason to doubt the competence of either speaker (2003, 93). At this point I invite the reader to consider whether this is a more accurate account of competent use of vague expressions, and to consider how they might react to disagreement or non-consensus about verdicts they themselves might make. Take the last shade you felt confident enough to say was red from the numbered series above — would you automatically question the competence of someone who disagreed with you about whether it was red? Or, indeed, would you question your own competence in the face of such disagreement?

If Wright's picture is more accurate than Bobzien's, then it seems that Bobzien ought to drop the claim that ideal speakers would not make conflicting assertions from the way she characterises them. But this creates a further problem. Recall that Bobzien also says that when ideal speakers make assertions, those assertions are always correct. She needs this claim in order to endorse the axiom

(T') $DA \rightarrow A$

given her interpretation of definiteness in terms of what ideal speakers 'can tell', and in turn her definition of 'can tell' in terms of what ideal speakers would be willing to assert. We can read (T′) as saying that if all ideal speakers would assert that some sentence A were true, then A is true, and it seems that this must be true if all ideal speakers only ever make correct assertions. But there's a tension if ideal speakers can make conflicting assertions *and* are always correct in the assertions they do make: if two such speakers *did* make conflicting assertions, both would be true, and so in turn Bobzien would in effect be saying that it's possible for contradictions to be true, which is presumably something she wants to avoid.

Bobzien (or her followers) are therefore faced with a choice. They could reject the claim that ideal speakers only make correct assertions, or they could insist that ideal speakers would never make conflicting assertions. The former is unappealing because rejecting this first claim undercuts the motivation for including (T′) in Bobzien's system of columnar logic — if ideal speakers can get things wrong, how could consensus from ideal speakers about a case entail that that consensus was correct? And rejecting (T′) is quite a loss; it's one of the few axioms of any logic for 'B' and 'D' that few theorists are willing to contest. Meanwhile the latter is unappealing for the reasons given above: it seems not to fit with the way speakers actually use vague terms.

The above considerations suggest that Bobzien's account of ideal linguistic behaviour is on shaky ground to begin with. However, I don't think any of this is conclusive, even if it does show the account to be quite seriously counterintuitive. We'll now go on to look at some more substantial criticisms.

3.2 Unclear Clear Cases and Clearly Borderline Cases

A key component of Bobzien's view, and columnar views more generally, is that all clear cases (of any expression) are clearly clear cases, and all borderline cases are borderline borderline cases. Keefe points out a discrepancy in how Bobzien's interpretations of 'clearly' (or 'definitely') and 'borderline' entail these two claims. Recall that

Bobzien's reason for saying that all borderline cases are borderline borderline was that if an ideal speaker cannot tell whether some expression applies to some object (and so if that object is a borderline case), they do not actively withhold judgement about that object, and cannot identify cases in which they are unable to make judgements because they cannot rule out that, in considering whether to make those judgements in the first place, they did not make a small error which might have caused them to assert either way. The discrepancy that Keefe (2015, 98) points to is this: why are ideal speakers this cautious about judging objects to be borderline cases, yet perfectly capable of ruling out any possibility of error for any clear case (since all clear cases are clearly clear)? Noting this discrepancy, and failing any justification from Bobzien, seems either to undermine axiom (4) or axiom (V) of the columnar view, since it suggests either that some clear cases might not warrant confident judgements from ideal speakers, or that some borderline cases might actually warrant that confidence. As a result, this objection threatens to undermine Bobzien's columnar view as a whole.

Now Bobzien could certainly come to the defence of axiom (4) by pointing out that the fact that clear cases are all clearly so is guaranteed by margins for error: an ideal speaker would only judge an object to be a clear case of some expression when they've ruled out the possibility of error (of a certain kind). But this would only serve to reinforce the possibility of clear borderline cases — ideal speakers could likewise be said to have ruled out the possibility of wrongly judging an object to be a borderline case, if they did so while leaving a margin for error, as Keefe points out (2015, 98-9).

Bobzien doesn't address this point directly, but offers a blanket reply to any objection which involves the possibility of identifying borderline cases. She says that if someone claims to have identified a borderline case, they must accidentally be using not the 'undecidable' sense of 'borderline case', which is relevant to vagueness, but instead the 'in-between' sense which I mentioned earlier, and which she takes not to be relevant to vagueness. Her reason for saying this is that, in identifying something as a borderline case, you've distinguished it from non-borderline cases, and so have identified it as 'in-between' the clear positive cases and the clear negative cases — that is, as a case which is neither a clear positive case nor a clear negative case (2015, 80-1). As such, you *can* tell its status, where you couldn't if it were an undecidable borderline case.

Keefe's response to Bobzien's worry is to say that distinguishing clear cases of 'red' and borderline cases of 'red' from one another does not entail that the borderline cases picked out are such that 'red' neither does nor does not apply to them; it only entails that it's not possible to tell either way. So she says that distinguishing these two kinds of case doesn't involve identifying any in-between borderline cases (2015, 94-5).

We can go one further than this, and say that Bobzien has implicitly presented us with a false dichotomy anyway, that either an object is an in-between borderline case or it is an undecidable borderline case. In some sense, the two are incompatible — Crookes Road can't be both an undecidable borderline case of 'steep' and an in-between borderline case with respect to 'steep' and 'shallow', because if it's an undecidable borderline case of 'steep', it's not possible to tell whether it's steep, and if it's an in-between borderline case, given Bobzien's notion of 'in-between borderline case', it's possible to tell that it's neither steep nor shallow, and so it's both possible to tell that it it isn't steep and not possible to tell whether it is. But there's no such contradiction in taking Crookes Road to be an undecidable borderline case of 'steep' and an in-between borderline case with respect to 'clearly steep' and 'clearly not steep'. Indeed, to identify an undecidable borderline case of 'steep' is *exactly* to identify an in-between borderline case with respect to 'clearly steep' and 'clearly not steep', since such a case is neither clearly steep nor clearly not steep. To identify a clear undecidable borderline case, then, is to identify both an undecidable and an in-between borderline case, just of two different sorts. It's therefore misleading to say that identifying such a case requires us to slip *from* one meaning of 'borderline case' *to* another.

4 Is Bobzien's Theory Inconsistent?

4.1 The Incoherence Objection

Let's suppose for the sake of argument that Bobzien has some satisfactory replies to the above considerations. She highlights another objection to her view, the 'incoherence' objection, and attributes the following version of this objection to Dever. A consequence of axiom (V) of the columnar view is that there are no clear borderline cases of any expression. But axiom (4) entails that if an object is *not* a borderline case of some expression, then it clearly isn't a borderline case of that expression — that is, any ideal speaker can tell that it isn't a borderline case. What's more, since ideal speakers lack no relevant information, if axiom (4) is correct, ideal speakers could tell that it was correct, and so could tell that they could tell any non-borderline case to be non-borderline. In which case, ideal speakers could identify all of the non-borderline cases in one of the orderings of objects we've been looking at. But, having done this, it seems that these speakers have identified all of the borderline cases as well — they could presumably assert that any object they *failed* to pick out in this way was not non-borderline after all, since they know that they can pick out any non-borderline case. And if this is so, there are clear borderline cases, contradicting axiom (V). Dever's argument therefore attempts to show that, on Bobzien's interpretation of 'clear' and 'borderline', the two key axioms of the columnar view are inconsistent (Bobzien, Draft).

Bobzien offers a 'refutation' of this objection by showing that no formal reconstruction of the argument is valid under the columnar view. Let's go through how we could formalise such an argument. First, we can formalise the claim that there are no clear borderline cases (of some predicate F) as $\neg\exists x_n(DBFx_n)$. This claim is a consequence of axiom (V). Next, we can represent the claim that all non-borderline cases are clearly not borderline cases as $\forall x_n(\neg BFx_n \to D\neg BFx_n)$. This claim is a consequence of axiom (4).[11] The claim that any ideal speaker can tell that all non-borderline cases are clearly non-borderline is just the definitised form of this last claim under Bobzien's interpretation: $D\forall x_n(\neg BFx_n \to D\neg BFx_n)$. This follows from its undefinitised form, since the undefinitised form is a theorem of the columnar view (that is, it follows from the axioms of the view), and any ideal speaker lacking no information could tell that it was a theorem (Bobzien, Draft).

Now, applying contraposition to $\forall x_n(\neg BFx_n \to D\neg BFx_n)$ gives $\forall x_n(\neg D\neg BFx_n \to \neg\neg BFx_n)$, or $\forall x_n(\neg D\neg BFx_n \to BFx_n)$: if an object isn't clearly a non-borderline case, it's a borderline case. Since this follows from $\forall x_n(\neg BFx_n \to D\neg BFx_n)$, and speakers can tell that $\forall x_n(\neg BFx_n \to D\neg BFx_n)$, $D\forall x_n(\neg D\neg BFx_n \to BFx_n)$ also ultimately follows from (4), since speakers can tell anything to be true that follows from things they can tell. So ideal speakers can tell that if an object isn't clearly a non-borderline case, then it's a borderline case (Bobzien, Draft).

So far the premises of Dever's argument are supported by the columnar view. But the last steps are where things become more difficult. It isn't enough that ideal speakers can't tell that some object isn't a borderline case to show that they can tell that that object *is* a borderline case. They first need to be able to tell that they can't tell that it's non-borderline, before making that inference. So $\forall x_n(D\neg D\neg BFx_n \to DBFx_n)$ needs to feature in Dever's argument. And this claim is a consequence of $D\forall x_n(\neg D\neg BFx_n \to BFx_n)$, as Bobzien points out. Using a form of the converse Barcan formula

$D\forall x_n Fx_n \to \forall x_n DFx_n$[12]

and the axiom

(K) $D(A \to B) \to (DA \to DB)$

in turn, we move from $D\forall x_n(\neg D\neg BFx_n \to BFx_n)$ to $\forall x_n D(\neg D\neg BFx_n \to BFx_n)$, and

[11] Bobzien proves this in her (Draft).

[12] Bobzien states it as written above, though it should presumably be $D\forall x_n\varphi x_n \to \forall x_n D\varphi x_n$.

then to $\forall x_n(D\neg D\neg BFx_n \rightarrow DBFx_n)$.[13] But how do we get from here to the claim that there are some clear borderline cases? Presumably we'd need $D\neg D\neg BFa_m$ to be true of some object a_m, so we could infer $DBFa_m$ from it. But Bobzien points out that $D\neg D\neg BFa_m$ is just not consistent with the columnar view, nor is it provable within it. We're trying to show that (4) and (V) are inconsistent, but we can't do that by assuming (4) and (V) to be true, and then just introducing an assumption which doesn't follow from either, but which is inconsistent with them both together. So, Bobzien says, Dever cannot have succeeded in showing the columnar view to be inconsistent (Bobzien, Draft).

Indeed, she goes on to point out that we can stop similar attempts to prove such an inconsistency by pointing out that her favoured logic, which contains the axioms (4) and (V), plus a few others, is provably consistent (Bobzien, Draft).

4.2 Another Kind of Incoherence

Is Bobzien's theory completely consistent, then? The axioms (4) and (V) are certainly mutually consistent, but (given some very modest assumptions!) her theory as a whole — rather than the logic she takes it to support — stands on shakier ground, as we'll now see.

Something seems correct about the 'incoherence' objection just presented, even if it ultimately doesn't succeed: ideal speakers could surely pick out at least some borderline cases of some expressions once they can recognise some doubt about the clarity of those cases. The question that Dever, for example, doesn't provide a satisfactory answer to is how they could do this. I'll now give an answer to this question that follows directly from Bobzien's columnar view.

[13] I'm not going to say anything about the converse Barcan formula, other than that it reflects the reasonable assumption that we're always dealing with at least the same number of objects in our orderings of objects.

Recall that Bobzien's margin for error constraints on ideal speakers guarantee that they will only assert that some expression applies to some object if all other ideal speakers would have the initial reaction that it did — this is supposed to stop speakers from making contradictory assertions. We can flip these constraints on their head: they likewise guarantee that if some ideal speaker would *not* have the initial reaction that some expression applied to some object, then no ideal speakers will assert that it does.

Now let's suppose that two ideal speakers were to look at the IC, and that one thinks (initially) that it's blue, while the other thinks it isn't. Then, as a consequence, some ideal speaker wouldn't have the initial reaction that it was blue, and another wouldn't have the reaction that it wasn't blue. In which case, given the margin for error constraints just described, no ideal speaker would assert that the IC was blue, and no ideal speaker would assert that it wasn't. So, when there's initial disagreement (or, at least, lack of initial agreement) about a case, Bobzien's margins for error imply that no idea speakers will make any assertion either way about that case. Under Bobzien's interpretation of 'borderline case', this means by definition that if there's initial disagreement about a case, it's a borderline case.

Now ideal speakers have all the information that would be useful in making judgements about cases, so if Bobzien's theory is correct, and some aspect of it is relevant, it doesn't seem unreasonable to suggest that they could tell that that aspect was correct. So let's assume that ideal speakers can tell that if there's initial disagreement about whether an expression applies to some object then it's a borderline case. Ideal speakers can do quite a lot with this piece of information: if they can identify two ideal speakers who have different initial reactions about whether *just one* expression applies to *just one* object, then they (and any other ideal speaker) can infer from this that the object they disagree on is a borderline case of that expression. So, on Bobzien's view, assuming that there's at least one case about which two ideal speakers would differ in initial reaction about whether an expression applies, there's at least one clear borderline case of that expression. The reasonable assumption that ideal speakers could initially disagree about some cases therefore generates inconsistencies within Bobzien's view.

Bobzien cannot reply to this argument in the same way that she responds to the 'incoherence' objection addressed in the previous section. This argument neither relies on a tacit assumption that there are clear borderline cases, nor does it try to show that Bobzien's *logic* is inconsistent. Rather, it tries to show that her interpretation of 'borderline case' in terms of ideal speakers' assertions is inconsistent with her proposed logic, given the modest background assumption that ideal speakers could differ in initial reaction about even one case.

4.3 A Possible Reply

What could Bobzien say in reply to this objection? One thing she might point out is that the way in which I've suggested that ideal speakers could identify borderline cases is a little roundabout: they first need to work out that variance in initial reaction to a case entails its borderline status (assuming Bobzien's theory is correct — an assumption I don't think she would reject(!)), then identify a case about which there is variance in initial reaction, and then infer that it is a borderline case.

Why would a process of investigation being roundabout discredit the results of that process? Bobzien thinks that ideal speakers should investigate cases in the most direct way: one natural thought is that, while an ideal speaker could reliably judge which shades count as red in an ordering by, for example, just parroting what other ideal speakers would assert, the kind of judgement we should be interested in is the kind that an ideal speaker arrives at by themselves. In judging whether a shade is red, they should actually look at it. The lesson Bobzien draws from this is that ideal speakers must investigate cases in the most direct way possible — typically involving 'perception, observation, intuition, imagination, calculation, measuring or inference' (2010, 7). If the process of investigation I've described isn't the most direct (or isn't sufficiently direct), she could reply that it's just not the kind of process that ideal speakers would make use of.

Now it's sensible enough to suggest that what speakers 'can tell' when we're thinking about vagueness is probably better construed as what they would assert on the basis of their own investigation, rather than merely what they would assert on testimony. But the process of judging an object to be a borderline case that I have described *is* perfectly direct. In Bobzien's sense, a borderline case of some expression is just one such that ideal speakers can't tell whether that expression applies to it. What information is relevant to deciding whether a shade is a borderline case of 'blue' in this sense? Under Bobzien's interpretation, presumably just looking at the shade won't be good enough — what would that tell someone about other speakers' hypothetical assertions about it? More relevant would be general information about ideal speakers and the patterns within which they make certain assertions: knowing these facts gets you closer to knowing which assertions ideal speakers will and will not make. And since ideal speakers' initial reactions to cases affect the assertions they make, finding out what some speakers' initial reactions to some cases would be seems to be a perfectly direct way of gathering relevant information. That's all that's involved in the process of investigation that I described in formulating the objection, and so the reply I would predict from Bobzien would not be successful.

5 Conclusion

Columnar views represent an attempt to preserve the phenomenon of higher-order vagueness in the face of a number of paradoxes. We've seen that these views do have the resources to get around those paradoxes. But Bobzien's version of the columnar view is independently unattractive for a number of different reasons — briefly, in relying on implausible characterisations of ideal speakers' linguistic behaviour and of borderline cases, as well as in suffering from a serious inconsistency.

I presented the columnar view as something of a last-ditch attempt to preserve higher-order vagueness, since columns and hierarchies are ultimately the only two theory choices available which have higher-order vagueness built in. This chapter should therefore indicate that we need to rethink the inclusion of higher-order vagueness in our theories, at least in the sense of 'borderline cases at all orders'. We'll now go on to look at how that can be done.

Chapter 4

(Re-)Characterising Vagueness

Right at the beginning of the thesis we started out with the question of how to characterise vagueness, and tentatively made a few assumptions, among them that an expression's vagueness involves it exhibiting borderline cases, being higher-order vague, and being epistemically tolerant. Now chapters 2 and 3 should have shown that the assumption of higher-order vagueness (understood as a claim about borderline cases) leads to inconsistency, or at least bad theories. In light of this, I'd now like to revisit the question of how to characterise vagueness. We'll take everything off the table and try to find a fresh approach to this question. We'll look at a few proposals (focusing in particular on two from Eklund and Horgan), seeing the shortcomings of each, and then drawing together the lessons we can learn from both to form a new proposal altogether. This proposal identifies a feature shared by all vague terms: if an expression is vague, there's a set of objects that could be ordered in a way that's relevant to that expression, about which competent speakers would not (between themselves) make a complete, stable set of judgements, even if they could investigate those objects in any possible way. We'll see in due course exactly what this means. After we've arrived at this feature, I'll discuss why I take this to give us only a necessary (and not sufficient) condition for vagueness, and gesture at some ways in which we could fill in the gaps, while noting that the necessary condition is the only one we need to make reference to in the remaining two chapters. I'll close by offering some discussion of some of the consequences of taking the condition I identify to be a necessary condition for vagueness, and how it compares to the other characterisations that we will have considered.

1 Tolerance

1.1 A Shift in Focus

In the previous two chapters we've been operating under the assumption that each borderline category corresponding to a vague expression (think 'borderline red', 'borderline borderline red', and so on) has at least one object in it. We got to this assumption by putting together two more basic assumptions, that each of the borderline categories just mentioned is itself vague, and that all vague expressions have borderline cases. Now if the assumption that each borderline category has some object that belongs to it cannot be made consistent in any plausible way — as I have shown — absent some other explanation, it seems that at least one of these two more basic assumptions must be false. So let's leave both of these assumptions to one side for now, and we'll re-evaluate them in due course. For now our goal is to arrive at a characterisation which doesn't automatically entail that vague terms exhibit higher-order (borderline) vagueness.

We also need to leave aside the assumption that vague terms are epistemically tolerant. This is because we saw an argument in chapter 1 from Greenough arguing that this assumption entails that vague terms all have borderline cases, an assumption we're trying not to commit ourselves to here. This argument is valid within classical logic, at least, so in the interests of theory neutrality we should discount the assumption that this argument starts from. Again, all of this will be re-evaluated later.

Having jettisoned all of these potential characteristics of vague terms, we don't seem to have much of what we started with left. We might still endorse some of the 'monotonicity' principles which we encountered in chapters 1 and 2. But these two features hardly illuminate the phenomenon of vagueness.

Let's see what else we can say. One thing to keep in mind is that all of the options we're going to look at reflect a shift in focus, away from questions about whether certain expressions apply to objects simpliciter, and instead focusing on questions about linguistic *behaviour* — when *speakers* would, or should, apply those expressions. I'm going to take these two questions to be about 'meaning' and 'use' respectively. So, for example, under this shift in focus, we're more interested in whether competent speakers would say that some object were, say, green, in some circumstances, rather than whether those objects *are* green, or whether it's true to say that they're green. We shouldn't read too much into the labels 'meaning' and 'use', but hopefully the idea is clear enough. Nothing turns on a particular divide between these two — and ultimately the characterisation we'll settle on won't make mention of either 'use' or 'meaning' — but we'll see that it's fruitful to focus on 'use' over 'meaning', broadly speaking.

1.2 Eklund's Characterisation

In chapter 1, we considered that the notion of 'tolerance' might feature in our characterisation of vagueness. We quickly ruled out the idea that vague terms might be characteristically tolerant — that is, it might be that small changes (of a relevant sort) cannot make the difference to whether a vague term applies or not — by noting that this would just rule out epistemicist theories of vagueness out of hand, by ruling out the possibility of vague expressions (unknowably) failing to apply after a small change of the right kind. Now if we wanted to focus on 'use' rather than 'meaning', and wanted to include something close to tolerance in our characterisation, probably the closest we could come would be to say that if speakers would assert that some vague expression applies to some object, they would also assert that it applies to a marginally different object. If they took Blake Street to be steep, they'd likewise take a road a fraction of a degree less steep to be steep too. But this suggestion isn't right. It's surely not plausible that speakers in general would allow themselves to actually be led, little by little, from some clearly green object, say, to a clearly blue one, happily asserting that each colour in-between (and so the clearly blue object itself) was green. At some point in this process, it's reasonable to stop asserting that successive shades are green. Note, though, that speakers can resist sorites reasoning of this sort while *also* refusing to draw a line between 'green' and 'not green' somewhere in the series. Other responses are available: they could go quiet when asked 'is this green?' of some object, or express uncertainty about some cases, for example, without thereby committing themselves to a snap change in judgement.

Eklund offers a proposal that gets a bit closer than this last suggestion, though.[1] His idea is that competent users of a vague term are disposed to *think* that it is tolerant. That is, they would be disposed to think that small changes of the relevant sort cannot affect whether that expression applies. This is distinct from the previous suggestion, he says, because this disposition is defeasible: the claim that vagueness involves tolerance is tempting — and Eklund claims that this is why the sorites paradox seems to rely on very plausible assumptions — but the paradox itself could give us reason to reject it in the end (2005, 41). Eklund's proposal is also distinct from that previous suggestion — that speakers' judgements aren't affected by small changes — because

[1] My comments here focus on his 2005 defence of this view, although he also defends this view in his 2002.

it's a claim about proto-theories of vagueness that speakers would be disposed to accept, rather than a claim about the 'first order' judgements that they'd make about particular cases: it involves a disposition to make a claim about 'green', rather than a disposition to make particular judgements about whether certain objects are green.[2]

One thing to note about Eklund's proposal is that it is a little abstracted from the use that speakers *actually* make of the relevant terms, since it makes a claim about what competent speakers *would* think about vague terms, rather than only what they actually do. While it would be nice to be able to give a characterisation of vagueness just in terms of actual language use, this slight abstraction is a good thing. For one, an account based on what competent speakers would do will not predict anything that would disagree with actual linguistic practice, since what competent speakers in fact do is what they would do. Such an account can also decide whether some terms are vague, even if there is not enough actual use associated with them to say either way on that basis alone. We can illustrate this with a term that Anna Mahtani describes, 'sprinkle', invented by her and her brother to describe sand that's no good for building sandcastles (2014). It may be that for some time no one considered the effect that small amounts of water could have on whether some sand was sprinkle or not, yet it seems that, on reflection, at least Mahtani herself can say with confidence that 'sprinkle' was vague when it was coined. Likewise, we might hope to make a ruling on whether 'child*' (the artificial predicate we saw in chapter 1) or '999,999th-order borderline case of 'steep'' are vague, even if these expressions don't have much use associated with them at all. In light of this, I'm going to assume from here that characterising vagueness in terms of what *would* count as competent use is acceptable, if we're interested in use at all.

[2]This is not to say that this disposition would have no connection to the judgements that speakers would make.

Returning to Eklund's proposal specifically, we might object that whether an expression is vague shouldn't depend on what (competent) speakers would think about them and how they work. In chapter 1, I envisaged a situation in which all English speakers were indoctrinated with the belief that vague expressions do in fact draw (albeit unknowable) sharp boundaries, pointing out that it seems that English would still contain vagueness in this scenario. Yet Eklund's characterisation of vagueness as presented so far implies that there would be no vagueness in that situation, since no one would be disposed to think that small changes could not make the difference to whether any expressions applied.

Eklund ultimately refines his characterisation in light of this kind of objection, saying that vague expressions are the ones which competent speakers would 'by default' assume to be tolerant — if they're competent with the relevant expression, they have the 'capacity to recognise that unless there are overriding considerations, the principle [that the relevant term is tolerant] should come out true' (2005, 50). To illustrate this idea: without indoctrination, the vast majority of the speakers I have in mind would presumably *not* form the belief that vague expressions aren't tolerant, so it seems that, given this revision, Eklund's characterisation correctly predicts that vagueness would exist in the post-apocalyptic situation I described.

1.3 No Sharp Boundaries in Use

Eklund's characterisation, taking vague expressions to be the ones which competent speakers would by default treat as tolerant, seems promising, but let's now contrast it with a related characterisation from Horgan. In the end we'll try to capture the positive aspects of both.

I'll first outline one part of Horgan's characterisation. We'll come to the full characterisation shortly, but before that I want to highlight why the rest is necessary by sticking close to the idea of tolerance to begin with. Horgan says that if an expression is vague, competent use of that expression involves not drawing any sharp boundaries. Put more specifically, Horgan says that it's a norm governing language use that if you assign one 'polar verdict' (such as 'true') to a sentence saying that some vague expression applies to some object, you should not assign an opposite polar verdict (such as 'not true' or 'false' in this case) to a sentence saying that that expression applies to a marginally different object (2010, 76). Now this doesn't quite amount to competent speakers taking vague terms to be tolerant by default (as Eklund thinks they do), though it isn't far off: even if they wouldn't necessarily think that they should continue to apply the relevant term after a small change, they won't actively make conflicting verdicts about two sufficiently similar cases.

2 (No) Overall Sets of Judgements

2.1 Some Test Cases

This much of Horgan's characterisation also seems promising as an answer to our 'characterisation' question, but it isn't the end of the story either. We're now going to look at a set of 'test cases' which show some details that the accounts we've seen so far either get wrong, or fail to address fully, and which give us a benchmark against which we can test other proposals.

The first case is the expression 'early thirties'. This case is raised in Smith (2008, 156) and Weatherson (2010, 81). The problem here is that someone *starts* being in their early thirties the moment they become thirty (which seems to be a precise matter), while the expression 'early thirties' is presumably vague — when does someone *stop* being in their early thirties? Yet Horgan's proposal (as stated so far) classifies it as precise, since competent speakers with enough information about people's ages will draw a sharp boundary between those to whom 'early thirties' applies and those to whom it does not on the basis of a small change in age.

Likewise, competent speakers would presumably not by default take 'early thirties' to be tolerant, and so Eklund's characterisation as stated so far seems to make the wrong prediction for 'early thirties'. Now Eklund does adapt his view in light of this kind of case to say that, for a term to be vague, speakers must be disposed to take it to be tolerant '*somewhere along its parameter of application*' (2005, 49, his emphasis). It's not immediately clear what exactly this amounts to — what exactly does a speaker need to believe to believe that a term is tolerant somewhere along its parameter of application? In the case of 'early thirties', for example, it can't be that making someone a day older doesn't affect whether they're in their early thirties, because making a 29 year old a day older can sometimes have an effect on that. Maybe, then, it's that making someone a day older can't take them from being in their early thirties to not being in their early thirties. But this isn't the same thing as saying that 'early thirties' is tolerant somewhere along its parameter of application — it's to specify exactly where along its parameter of application it is tolerant.

Perhaps a way to spell out the claim that a term is tolerant somewhere along its parameter of application would be to say that a certain *disjunction* is true: in the case of 'early thirties', for example, we could say that it's tolerant in this respect if *either* making someone a day older can't take them from being in their early thirties to not being in their early thirties *or* making someone a day older can't take them from *not* being in their early thirties *to* being in their early thirties. In this case, the first disjunct is the one that Eklund claims to be the one that competent speakers would by default take to be true. This seems to work by Eklund's lights, though in the next section we'll see a way in which we can eliminate the need to use a disjunction by appealing to something more general.

The second test case is the artificial predicate ‘child*’. Recall that someone is a child* if they’re younger than 17, and not a child* if they’re older than 18, and so by stipulation it’s indeterminate whether those who are between 17 and 18 are children*. Eklund’s characterisation seems to get this one correct: competent speakers wouldn’t by default take ‘child*’ to be tolerant (since it’s at best *unclear* whether someone who turns 17 stops being a child* at that moment), and so it comes out as precise. It’s less clear what Horgan’s view (as stated, at least) says about this case; really we need to know what counts as a ‘polar opposite’ verdict before we can say whether competent speakers would apply polar opposite verdicts to marginally different cases when it comes to ‘child*’. It seems inappropriate to say, for example, that the sentences ‘someone who’s a moment under 17 is a child*’ and ‘someone who’s exactly 17 is a child*’ go from true to false, or perhaps even from true to not true, since the status of the latter is by stipulation indeterminate. Still, it *would* be appropriate for someone to assert that between these two sentences there is a clear change from determinacy of status to indeterminacy. If Horgan’s notion of a ‘polar verdict’ allows for this, then Horgan’s characterisation correctly predicts that ‘child*’ is precise; it’s just not clear whether it does. In any case, whatever account we arrive at in the end, we should make sure that it predicts that ‘child*’ is precise.

The third test case is actually a set of cases. We need to ask: what does the correct account say about different borderline categories? Supposing that ‘green’ is vague, does the account predict that ‘borderline green’ is vague? ‘Borderline borderline green’? How far does this go, if at all? I raise this last set of cases not necessarily as a set of cases about which either Eklund or Horgan gets things wrong; rather I raise it as something which any account needs to have a defensible answer to.

2.2 A Broader Characterisation

We can do better than either of Horgan or Eklund's proposals. A good starting point for doing this will be to introduce the other half of Horgan's characterisation — this is going to reveal an important element for us to adapt. Horgan says that, in addition to not making opposite polar verdicts about sufficiently similar cases, competent users of a vague term won't 'affirm any determinate overall assignment of statuses to [sentences each saying of a member of a gradually changing set of objects that the relevant expression applies to it]' (2010, 76). This is about as much detail as we get from Horgan, so let's use this as a 'jumping off' point and try to extract what this suggestion gets right and fill in some details, whether this fits with Horgan's intentions or not.

What's good about Horgan's characterisation, filled in this way? It captures the idea that vagueness isn't just about tolerance (or some phenomenon related to it), but rather than vague terms exhibit some *overall indeterminacy* when it comes to sets of small changes. Informally speaking, it's somehow difficult to completely pin down vague terms. Including this kind of thought in our characterisation will help us to deal with some of our problem cases. Horgan's proposal also makes a claim about competent use of vague terms rather than their meaning. Here this is helpful since it means that the characterisation doesn't straightforwardly rule out epistemicism — consider its 'meaning' counterpart, that there's no overall fact of the matter about which objects vague terms apply to.

How does this deal with the first two test cases that I gave — 'early thirties' and 'child*'? With 'early thirties', while it seems that speakers could happily draw a line between those younger than their early thirties and those in their early thirties, it seems that they won't (broadly speaking) be able to say exactly which of a group of people, going from 29 to 40 in small increments, are in their early thirties and which aren't. On this score, 'early thirties' can be made sense of as vague. A consequence of this, though, is that Horgan's first condition for vagueness (that competent speakers must not change polar verdicts on the basis of a small change) must actually be false, since 'early thirties' still fails to meet this condition.[3] We should therefore think of Horgan's view from here as just being this second condition of no overall assignment of statuses by speakers.

It's as unclear as before how Horgan's account (so modified) deals with 'child*'. If 'is determinately true' and 'is not determinately true' are 'statuses' that you can assign to the relevant sentences saying of certain people that they are children*, then Horgan's account will correctly predict that 'child*' is precise, since there is a determinate overall set of statuses that a speaker could in principle make to a set of sentences about whether certain people are children*. If not, the theory doesn't rule 'child*' out as being precise.

We can clear up this ambiguity. To see how, we need first to deal with another problem for Horgan's account. Don't speakers sometimes draw boundaries in their use of vague terms, and likewise don't they completely categorise some cases even when using vague terms? For one, it seems perfectly acceptable to make stipulations for certain purposes — a pot of tea isn't 'brewed' after exactly three minutes of making it, but we might just agree to draw the line there and pour it out then, for example. What's more, and perhaps more controversially, in talking to a few people about vagueness, I've presented them with some set of objects that's supposed to be exemplify the phenomenon (in fact a piece of paper with shades going from red to

[3]Could Horgan apply a similar disjunctive strategy to the one I suggested when talking about Eklund's claim that speakers must take vague terms to be tolerant 'somewhere along its parameter of application' here? It's harder to see how, since a lot of things may be able to count as 'polar verdicts' — it's not limited to speakers saying either a sentence is true or saying it's false. So we can't just spell Horgan's condition out in the case of 'early thirties' as 'competent speakers will either not go from a positive polar verdict on whether someone is in their early thirties to a negative polar verdict on the basis of making that person a day older, or they will not go from a negative polar verdict on whether someone is in their early thirties to a positive polar verdict on the basis of making that person a day older': these don't exhaust all of the options.

orange) and have been met by insistence that *this* or *that* shade is the last red one.[4] Assuming 'red' is in fact vague, Horgan's characterisation (modified or not) judges such people, as well as those who make the kinds of stipulation just described, to not be using 'red' competently. This would be an interesting result if it were correct! It's also worth noting here, thinking about getting a better account out of this kind of consideration, that in a trivial sense, speakers can always draw boundaries and give complete sets of responses to cases when considering gradual changes to objects: they can draw boundaries between the shades they know to be red and the ones they *don't* know to be red, for example, shades which they'd be happy to call red and shades which they're not so sure about, and shades about which they can say something and shades about which they refuse to say anything. These are not all necessarily 'polar opposite' verdicts in Horgan's sense, though.

How could Horgan makes sense of the first two kinds of boundary drawing that I just mentioned? Thinking first about speakers making stipulations, it seems that these speakers should not be classed as incompetent users of vague terms. Hopefully that much is obvious. So Horgan would need to modify his characterisation of vague terms further to make room for this use. One way to do this would be to introduce a notion of 'sincerity', saying that stipulating the boundary between 'brewed' and 'not brewed' to be here or there in some sense 'doesn't count' as a relevant assertion because it's in some sense not sincere — it's used to draw a boundary, but isn't supposed to represent where a boundary *really* is. A modified version of Horgan's characterisation could therefore say, in effect, that speakers won't *sincerely* assign a status to each sentence in a series saying of gradually changing objects that the relevant expression applies to them. This strategy isn't ideal, though, because it's not clear that stipulations aren't

[4]Note that each of these examples shows someone drawing a line in a way which also serves to give a complete assignment of statuses to cases. Talking about stipulations in this context therefore has the same implications whether we're thinking about vagueness as 'no sharp changes of verdict' or as 'no overall assignment of statuses'.

made sincerely. The lines that speakers draw when they stipulate don't need to be 'pretend lines'; they just need to be the ones that apply in a particular context.

Before coming to what I take to be a better way to deal with this problem, let's think about our second problem case: people who claim to have separated all of the red shades from the non-red shades over a gradual transition between the two. One thing we might consider saying would be that this use is not competent, since it more or less involves a speaker denying that the relevant expression (in this case, 'red') is vague, when it clearly is. Horgan's view (in its modified form as well) seems to imply this response, although he doesn't explicitly acknowledge it. It's also an implication of Eklund's view, and he says that someone who doesn't feel the force of the sorites paradox because they can't see vague predicates as tolerant is missing something important about those predicates (2005, 41-2). But I don't think this is correct. My intuition is that these people can clearly also be competent users of the relevant terms. They may have 'missed something' about vagueness, but it seems that it shouldn't be a condition on competence with an expression that a speaker *recognises* every feature of that expression. On the other hand, I concede that not everyone shares this intuition (Eklund doesn't, for example), and so it would perhaps be better if our final characterisation of vagueness didn't force a judgement on this question. With that in mind, I'd like to look now at a feature we could appeal to in addressing both problem cases.

What should we say about stipulated line drawing and those who insist they've drawn lines in the correct places? That they show that vagueness isn't characterised by competent speakers coming to no overall set of verdicts about cases, but rather in their coming to no *stable* set of verdicts. The idea here is that if speakers *do* draw lines (and, by extension, form a verdict on all of the relevant cases), other competent users of the relevant expressions, as well as themselves at other times, would not necessarily agree with them about where they drew those lines, except for the sake of stipulation in some particular context. So, for example, let's say that you and a friend are putting up a shelf. Between you, you decide that the two inch screws that you've got are the 'long' ones, and that anything shorter isn't long. Now it seems inappropriate, at least, to expect *someone else* to understand that you're asking for a screw of two inches or more if you ask them for a long screw, and what's more, you and your friend might draw the line somewhere else for your next DIY project. Likewise, while Ahmad (not to name names) thinks he's found the last shade of red on the paper I showed him, other speakers won't necessarily agree with him on where he has drawn the line between 'red' and 'not red' — in fact, I suspect he would be in a minority of people who would draw such a line in the first place.[5]

We could fit this notion of 'stability' into our already modified version of Horgan's characterisation of vagueness by saying that if an expression is vague, competent users of that expression won't, in a stable way, assign statuses to all of a set of sentences each saying that it applies to a member of a gradually changing set of objects: if a speaker were to make complete assignments of statuses, that set of assignments would at least not be widely shared by others, or even necessarily by that person at a later time.

Let's now return to an ambiguity in Horgan's picture that I mentioned before. What exactly counts as a 'status' or a 'polar verdict' on a sentence? I don't think it's going to be helpful to address this question directly. Instead, I'd like to do so indirectly, diverging from Horgan's framework at this point, suggesting instead the following (provisional) characterisation: if an expression is vague, there is a (possible) set of objects that represent gradual changes with respect to that expression, such that competent speakers would not, in a stable way, categorise them into those to which that expression applies, those to which it does not, and any others. Conversely, if compet-

[5]It doesn't matter whether this suspicion is correct or not. However, it does seem to be backed up by the results of a study conducted by Raffman et al (see chapter 5 of Raffman 2014), as well as the results of one conducted by Alxatib and Pelletier (see Alxatib and Pelletier 2011), though I'm not going to offer any comment on these sources.

ent speakers can between themselves sort any objects into these three categories in a stable way the expression in question is precise. 'Other' here is to be interpreted very broadly: any response a speaker could give that wouldn't amount to them having asserted that the relevant expression applied or did not is counted. This interpretation will allow us to get clear on more cases than Horgan's characterisation, since it leaves little room for ambiguity.

We just need to add one more element to this characterisation. Presumably the expression 'younger than thirty' is not vague, since being thirty is a precise matter (or could at least be plausibly construed as a precise matter). Yet the characterisation we've got so far might fail to rule this out as precise. Suppose speakers were presented with gradually younger people, starting with a forty-year-old and ending with a twenty-year-old, say, and were asked whether each was younger than thirty, without being given their exact age. It seems likely that speakers would judge at least some to be younger than thirty, and others not to be, but that no stable overall categorisation would emerge from these speakers' responses. In which case, 'under thirty' would not be just as precise after all under the characterisation we're considering. But it could easily have been taken not to be precise in different circumstances: those where the speakers in question had crucial relevant information — in this case, information about people's ages. Part of why 'under thirty' is precise, then, is because it's in principle possible for speakers to get information that allows them to arrive at a stable categorisation. In contrast to this, there seems to be no way even in principle for speakers to arrive at a stable answer to the question of which objects 'red' applies to, which it doesn't, and which warrant some other response. Our characterisation should therefore only treat as relevant linguistic behaviour that's not a consequence of not having carried out an in principle possible investigation.

I don't intend to give a full answer here to the question of what counts as an 'in principle possible investigation'. But, to give some illustration about what I have in mind here, let's look at the kinds of claim made about *resistance to enquiry* by Sorensen that I brought up in chapter 1. Sorensen likens questions such as 'where do you draw the line?' (or in our case, lines) to questions such as 'does a glass house have any windows?', and says that even a persistent enquirer will 'hit a wall' (2001, 24). You can find out everything that's relevant to answering these questions, and still find yourself stumped. Now, this claim doesn't fit smoothly with my claim that competent speakers *will* sometimes make a judgement about all of a relevant set of cases, nor the claim that competent speakers will disagree on some cases, even if they have all the same information available: why do some speakers hit walls in certain places and others don't, if an answer to the question of which objects 'red' applies to, which it doesn't, and which warrant some other response, is not possible to give? The point to make here is that it's not that it's not possible to give such an answer, but that it's not possible to reach a *stable* answer: the question of which judgements to make about our set of shades is one that seems to resist enquiry *between* even competent speakers — between them, they will hit a wall because they will all, even as persistent enquirers, fail to reach a stable answer.

Let's draw out our characterisation as it stands, then: we'll say that if an expression is vague, there could be a set of objects gradually changing with respect to that expression such that competent users of that expression wouldn't make judgements which stably categorised those objects into those to which the expression applies, those to which it does not, and any others, even after exhausting any investigation that they could carry out.

2.3 Necessary, but not Sufficient?

Let's suppose that the characterisation as I've stated it so far points to a feature shared by all vague terms. That is, let's suppose that we've correctly stated a necessary condition for an expression's vagueness (we'll check this in more detail in a moment, but suppose for now). Is it also a sufficient condition for vagueness? There are a few kinds of case that we can point to which cast doubt on that, which satisfy this condition and yet this doesn't give us enough reason by itself to think that they are vague (whether or not they actually are).

First, consider other kinds of context-sensitive terms. Is there a stable way in which speakers would judge objects on whether an *ambiguous* term such as 'bulb' applies to them? Perhaps not, as the ambiuguity between 'plant organ' and 'light source' makes it difficult to say whether something is a 'bulb' simpliciter — it depends on what you mean by 'bulb'. We built into our characterisation so far that for vagueness there need to be no stable judgements *about some set of gradually changing objects*, but for ambiguous terms like 'bulb' this instability could well crop up even when looking at a set of gradually changing objects — say, a set of cells gradually developing into a tulip bulb — because it could remain unclear which sense of 'bulb' is intended, and so some speakers could fail to give a clear verdict when considering something that's clearly a tulip bulb where others do give a clear answer. But that shouldn't entail that 'bulb' is vague; ambiguity of this sort isn't a kind of vagueness.

A related example is to do with people who use terms to mean closely related, but slightly different, things. For example, American English use of the word 'professor' is more liberal about who counts than its British English counterpart: in American universities it can refer to any teacher, where in British universities it only refers to teachers of a certain rank. This means that there will be no way for speakers to categorise people into those who are professors, and those who are not (and any others), since they will inevitably apply the term differently to some cases in a widespread way. But, whether or not 'professor' is vague, this presumably shouldn't *entail* that it is vague; this seems to be a display of quite a different phenomenon.

Second, consider evaluative moral and aesthetic terms. People are widely (and resolutely) divided on what they consider to be 'morally justified' actions, and likewise what they consider to be 'good' songs or 'funny' jokes. As a result, if the characterisation I've offered is taken as giving a sufficient condition for vagueness, it seems that it automatically judges all of these terms to be vague. Again, this problem arises even with the stipulation that there has to be no stable set of judgements about a particular set of gradually changing objects. Consider two people, one of whom thinks that lying is never acceptable, and one who thinks that it is acceptable in some cases. We might present both people with scenarios in which someone lies, but each time for a slightly greater benefit. It seems that, whatever these two people's judgements about these scenarios are, 'acceptable action' would come out as vague, since there'll be no broad agreement on the cases (even if there was agreement on some cases being unacceptable). It may well be that 'acceptable action' is vague, but presumably this set of conflicting responses shouldn't demonstrate that.

These examples all seem to show that there's an aspect to vagueness that's not captured by the characterisation that I've put forward so far, and it's not obvious exactly what it is in each case. For the purposes of this thesis, this isn't a big problem: in the next two chapters I'm not going to make use of any claims stronger than that I've identified a feature shared by all vague terms. It's also worth pointing out that this doesn't mean that this chapter hasn't moved the discussion about characterising vagueness forward. For one, the necessary condition we've arrived at does better than Eklund's in giving a characterisation which captures the sense in which vagueness involves some 'overall indeterminacy' surrounding gradually changing objects (and thereby removing the need to invoke a disjunction in making sense of 'tolerance along some parameter of application'), and in not forcing an answer to the question of whether speakers who 'draw lines' are competent users. And we've added some important details in forming a characterisation similar to Horgan's. It's also not clear how either

Eklund or Horgan's characterisations deal with the three cases I've just described — in Eklund's case, for example, it's not obvious what the default attitude a competent speaker should take to whether 'acceptable lie' is tolerant or not, as it's going to depend on whether you think there are any acceptable lies in the first place; Horgan is in just the same situation as my characterisation.

Is there more we can say about how to fill the gaps left by the cases I've just raised? It would be easy enough to make room for the first kind involving ambiguity — 'bulb' and 'bulb' — we could make our condition that a term is vague when there's no stable overall assignment of 'applies', 'doesn't apply' and 'other', *after disambiguation*. That is, we could say that, once all (hypothetical) speakers are agreed on one sense of the relevant term (say, 'plant organ' in the 'bulb' case), if they don't make a stable overall assignment of that kind, then the term is vague. This strategy could also be made to work for the 'professor' case — we could say that the (at least two) meanings of the word are disambiguations. It's not clear how resilient this strategy is by itself, though — are two people who disagree about what counts as 'art', for example, necessarily just deploying two different concepts? That is, in order to make sense of the disagreement that speakers might have about art that's clearly not related to vagueness, do we need to say that those speakers are working from different ways of disambiguating the term 'art'? And if so, would these speakers be talking past one another? It would be better if our characterisation didn't require us to settle these questions.

Really what we want to capture is that when there's vagueness, there are unstable sets of judgements that in some way centre specifically around a set of gradual small changes. To illustrate how we might do this, let's compare two ways in which we could fill in the details of my 'acceptable lie' example from earlier. In the first, we'll say that one speaker would say that no lies are acceptable, and that another speaker thinks that some lies are acceptable, and can consistently pick out exactly which lies they think are acceptable and which they think aren't. In the second, we'll again say that one speaker would say that no lies are acceptable, and that the other speaker thinks that some lies are acceptable, but this time we'll say that this second speaker draws the lines between judging lies to be acceptable, not acceptable, and giving some other response, in different places at different times. In the second of these two cases, we seem to have a better claim to saying that 'acceptable lie' is vague, and it's the inconsistency across time of that second speaker's responses that seems to make this so.

We therefore want to make sure that an expression isn't predicted to be vague when there's stable disagreement about cases: what's needed for vagueness is *unstable disagreement*; some speakers need to fail to reliably carve out a set of cases in any particular way. So we might replace the necessary condition we have so far with the following condition: if an expression is vague, for each of some portion of the competent users of that expression, there's a (possible) set of gradually changing objects (relevant to that expression) about which they would not reliably give any particular set of responses. This would entail that all vague terms satisfy the necessary condition we've been working with so far: when the new condition is satisfied, competent users will not between them make stable sets of judgements about the relevant sets of gradually changing objects; there will always be the possibility that one such user would make a different set of judgements at a different time.

This kind of condition seems like a promising route to a set of necessary and sufficient conditions to me, but it opens up some questions. For one, what proportion of competent speakers would need to fail to reliably produce responses in this way? Could just one speaker making judgements in this way, even in the face of a broad consensus on cases from other speakers, be enough to make a term vague, for example? We might also ask, if vagueness requires that for each of this set of speakers (however big it is) there's a (possible) set of cases about which they wouldn't reliably make judgements, should we also say that this set of objects needs to be the same one for all of those speakers in order for there to be vagueness? Or is it enough that each such speaker has a (possible) set of cases about which their judgements would be unreliable? I leave these questions open — for my purposes the necessary condition that I gave before is really the important detail — and from here I'll leave to one side this extra condition I've just outlined.

2.4 Testing the Characterisation

We'll conclude this chapter by seeing how well the characterisation that I've given deals with the test cases that have come up over the course of this chapter. Before we start, it's useful to notice that, since we only have a necessary condition for vagueness so far, the most our characterisation can be asked to do is to correctly predict as many precise expressions as possible to be precise, and to correctly identify any vague expression as satisfying a necessary condition for vagueness — we don't have the resources to outright judge any expression to be vague given this characterisation alone.

The characterisation I've offered won't predict that 'early thirties' is precise: although it will say that we could expect stable lines to be drawn by speakers between those to whom it does apply and those to whom it doesn't (when looking at people just over thirty and just under), the fact that there are no such stable lines when looking at older people still means that 'early thirties' satisfies the necessary condition given. This isn't enough to predict that it's vague, but that isn't the aim of the characterisation as it stands.

In contrast, it *will* predict that 'child*' is precise, since there's a stable way of dividing up all cases into children*, non-children* and any others. There could well be a diversity of answers among speakers about the status of people who are between

17 and 18, but this is accounted for by how broadly 'other' is interpreted. It therefore seems that this characterisation makes the correct prediction.

What about borderline categories? When we have a vague term like 'steep', is 'nth-order borderline case of 'steep'' vague for any (or indeed all) n? It's difficult to answer these questions directly because the notion of a 'borderline case' is ambiguous, and, as I pointed out earlier, ambiguity can interfere with speakers' judgements in a misleading way. So let's try to refine the question. Is 'nth-order *epistemic* borderline case of 'steep'' vague for any n? In our terms, this amounts to the following question: would competent users of 'can be known to be steep' (and/or 'can be known to be knowable to be steep' and/or 'can be known to be knowable to be knowable to be steep', and so on) make stable sets of judgements about successively steeper objects, which divided them into those ones which could be known to be steep (or known to be knowable to be steep, and so on), those which cannot, and any others? It seems unlikely for any of the terms in question: each of them seem clearly applicable to some cases (Blake Street, for one) and clearly not to others (West Street, say), yet it's hard to see where speakers could identify a point at which they all felt too uncomfortable to outright judge cases to be knowable (or knowable to be knowable, or…). So 'nth-order epistemic borderline case of 'steep'' looks like it satisfies our necessary condition for vagueness for any n we like. We get the same result for semantic interpretations of 'borderline case', too — consider what happens when you replace 'can be known to be' and 'knowable to be' with 'can truly be said to be' and 'truly described by', for example.

An interesting observation we can make on this theme is that, for any vague term, speakers will never identify a stable 'other' category in the ordering that's relevant to a vague term. This is trivial in the case where no speakers identify any objects as 'others'. When some speakers *would* identify some objects as 'others', if they were to stably identify some particular set of objects as 'others', this would also amount to them stably picking out all the objects to which that term applied, and all of those to which it didn't, and so the term wouldn't be vague after all. This seems to explain why vague terms exhibit the appearance of 'higher-order vagueness' described in the previous paragraph: for the apparent higher-order vagueness in this sense to be 'cut off' at some order, there would need to be a stable 'other' category drawn by users of vague terms.

It's worth noting that we *haven't* just demonstrated that vague terms also all satisfy a necessary condition for exhibiting higher-order *borderline* vagueness. In chapter 1 I distinguished between this sense of higher-order vagueness and higher-order vagueness in the sense of the apparatus we use to describe vague terms (and so the term 'borderline case of' and its iterations) itself being vague. Since we're not endorsing the assumption that all vague terms have borderline cases, even if 'nth-order borderline case of 'steep'' is vague for all n, this doesn't allow us to infer that 'steep' has nth-order borderline cases for all n, and so we need only have the necessary conditions satisfied for the second, weaker, sort of higher-order vagueness.

Let's finish with a word about theory neutrality. The shift in focus that we've been making use of (from 'meaning' to 'use') allows us to keep things theory neutral. From an epistemicist perspective, the 'meaning' counterpart that we might construct from the 'use' characterisation that we've arrived at, (presumably) that there's no *correct* way to sort objects into those to which a vague expression applies, doesn't apply, or has some other relation, is a non-starter. But the characterisation we've arrived at does better than this in that it's *consistent* with epistemicist, as well as non-epistemicist, views, while not *entailing* either.

3 Conclusion

This chapter has extracted some key elements from some promising characterisations of vagueness (focusing on Eklund and Horgan's in particular), and has improved on some of the flaws of each, culminating in a new necesary condition for vagueness. I've shown why I don't take this to be a sufficient condition (while gesturing towards a way to get one), but have shown how the necessary condition itself deals with the problem cases presented in this chapter. In the next chapter, we'll see that this characterisation, while theory neutral, can lead to surprising results.

Chapter 5

Is Vagueness Impossible? Could There Be a Correct Semantics for Vagueness?

In the previous chapter I gave a necessary condition for vagueness that I said that any theorist about vagueness could, and should, accept: in short, that competent users of a vague expression couldn't, between them, clearly categorise some set of objects with respect to that expression. This chapter will explore the significance of this condition, showing that any semantics for vague terms which claims to accurately reflect this feature of those terms' use must inadvertently fail to do so, either explicitly or implicitly.

I'll start with a discussion of what it would take for the semantics of vague terms to 'match' the way they are used. This discussion will leave us with some resources that Fine exploits in order to offer a proof that we can construe as showing that any 'correct' semantics of vagueness render it contradictory. I'll explain this proof in some detail, and show that it can be offered using more limited resources even than Fine suggests, although I'll suggest that it isn't conclusive, since it uses a rule of inference (reductio) that is contested by some theories of vagueness.

I'll then offer my own argument for a similar conclusion, using Horgan's 'transvaluationist' theory of vagueness as a means to demonstrate how this argument could be applied to specific theories of vagueness.

1 How Vague Terms are Used and What they 'Mean'

Let's restate the condition for vagueness that we arrived at in the last chapter. We said that, for any vague term, there's a set of objects, gradually changing with respect to that term, such that competent speakers would not, consistently with each other or themselves over time, make judgements on them that neatly divide them into those which they take that term to apply to, those which they take it not to apply to, and any others, even when they've exhausted all relevant lines of enquiry. If this is how vague terms are used, what can we say about their 'meanings' (in the sense I used this term in the last chapter) in light of this? In particular, what bearing (if any) does this have on the question: for any vague term, and its corresponding set of gradually changing objects, is there a *correct* way of categorising them into those to which it applies, those to which it does not apply, and any others?

It seems that there can only be two answers to this question: 'yes', or 'no'. Now, any 'yes' answer would not require the person giving it to be able to give such a categorisation for any vague term — though being able to do so would certainly help if it were possible! — and indeed we should not expect to find any such categorisation, since by our assumption competent speakers can't achieve this, even when allowed to investigate in any way they like. Still, the answer 'yes' gives rise to a number of different questions. It would obviously take too much space to look at even a significant number of these in any depth, since 'yes' is precisely the answer given by epistemicists (and there are many ways to be an 'epistemicist'), for example, and a significant amount of ink has been spilled over that view alone. However, I'd like to focus for a moment on a particular kind of objection to any view fitting this description that specifically relates to the necessary condition for vagueness we're working with.

One version of the objection is this: suppose that there is some way of categorising all of the relevant objects relative to some particular vague expression. Why would it be that no one can find out what the correct way to categorise those objects is? To give one response, Williamson argues that it would not be humanly possible to get to grips with all of the ways in which speakers in fact use vague terms, and so it's likewise no surprise that it's not possible to arrive at (or even imagine) any clear answer about what a correct categorisation of the relevant objects would look like (1997, 220-2). To know the correct way of categorising a set of shades going from red to orange, for example, we'd have to know exactly how speakers use the terms 'red' and 'orange' over a variety of times and across a variety of contexts, and aggregate all of this information appropriately. But he takes both of these tasks — collecting or making sense of such a huge amount of linguistic data — to be impossible in practice; as he puts it, '[meaning] may supervene on use in an unsurveyably chaotic way' (1994, 209).

Even if this is a suitable reply to the question 'why could no one find out the correct way to categorise objects?', Horgan offers a slightly different objection that this kind of reply does not seem to give an answer to (2010, 86-7). Let's grant that it would not be feasible to collect all of the linguistic data that we'd need in order to find out the correct categorisation of the gradually changing set of objects relevant to any given vague expression (if there is such a categorisation). If it *were* feasible, what could this information tell us if we had it? We've already agreed that the use of any particular vague term doesn't allow for any particular categorisations of objects to be straightforwardly read from it, so, given the sum of such use, how could any such categorisation be a 'result' of it? What would explain the actual lines being drawn in one place rather than another? Are there some additional principles governing how overall use interacts with meaning this way, and how it fills in the gaps?

Here's a way to illustrate this point. How did the computer that I'm typing this on get made? The answer to this question is presumably extremely complicated, involving corporations with huge numbers of people collecting resources, manufacturing parts, buying and selling those parts, and finally assembling them, not to mention all of the work that went into working out how to make such a machine in the first place. Certainly it would be difficult to get a grip on each stage in this process, even just as it relates to *this* computer. But still, we can imagine at least *in principle* how all of these events might have come together. On the other hand, can we in principle imagine how the complex patterns of use surrounding vague terms could give way to a kind of precision at the level of meaning, in any given context? What about use would explain that precision?

From one epistemicist perspective, this kind of objection gets things the wrong way around. The thought goes (in, for example Williamson (1997, 217-8)) that, vagueness aside, there are very strong reasons to think that classical logic holds, and so, when we do come to vagueness, the stance we should take ought to be 'classical': there must be complete, correct ways, of categorising all of the cases. From this perspective, the challenge of explaining how, and where, the meanings of vague terms draw their lines, perhaps leaves more room for agnosticism — once you know for sure that the lines are drawn somewhere, does it matter exactly how they get there? It's perhaps possible to give the broad outlines of how (Williamson, for example, says that ultimately use must determine where these lines are), but that's as far as we can go, and this is just Williamson's point when he says that meaning may supervene on use in an unsurveyably chaotic way (1994, 209).

Now, I don't intend to engage with Williamson's reasons for thinking that classical logic is correct. However, I will point out that, given where this discussion is going to lead, I will argue that we can endorse classical logic *even if* we accept my assumption that, in a correct semantics for vague terms, their meanings would reflect no complete categorisations of cases into 'applies', 'doesn't apply', and 'other'. And if this argument gets things right, the reasons cited by Williamson for thinking that the meanings of vague terms draw these categories is undercut: endorsing classical logic now doesn't force there to be sharply drawn categories, and so taking on the assumption that these categories are drawn becomes much more suspect — as Keefe puts it, whatever explanation is given for how these lines are drawn (or whichever explanation is gestured towards), 'this does not touch on the implausibility of the claim that such lines *are* drawn' (2000, 80).

I don't raise these objections to once and for all show that the answer to the question 'is there a correct way of categorising the set of gradually changing objects that's relevant to any particular vague term?' is 'no'. Rather, I want to provide some strong reasons, using the condition for vagueness that we arrived at in the previous chapter, to move forward on the assumption that the answer is 'no'. Strictly speaking, then, epistemicism is still a 'live' option, since I can't conclusively rule this out — and it's certainly not *ruled out* by the characterisation of vagueness that we arrived at in the previous chapter — but we'll go forward on the assumption that there are no correct categorisations (of the sort I've described) of the relevant set of gradually changing objects relative to any vague term.

2 Fine's 'Impossibility' Theorem

What's the significance of the assumption that I've just motivated? Fine (2008) has made striking use of it (or, at least, an implication of it) in order to offer a proof of the claim that vagueness is *impossible*. We can interpret this claim in a number of different ways: to mean that there are no vague predicates, or that the conditions for being a vague predicate are mutually inconsistent, to give some examples.

2.1 An Outline of the Argument

Before we get down to the finer details of the proof, let's go through an outline of how it works and some of the ideas that are driving it. Take a set of sentences saying of each of, say, a set of shades going from red to orange that it is red. Fine takes there to be two basic things that we should be able to say in assigning semantic statuses to these sentences (where 'status' includes their being true or false, and their being definite or borderline cases at particular orders). The first is that we should be able to assign these statuses in a way which is consistent with us saying that it's true that the reddest shade is red and that it's false that the least red shade is red. The second is that we should be able to assign statuses in a way which is *not* consistent with us assigning a 'sharp' set of 'responses': whichever statuses we think the sentences have, what we say about them shouldn't be consistent with us taking each of them to have a semantic status such that any other sentence has either the same status or an opposite (that is, inconsistent) status (think true/false, definite/not definite, for example). This second requirement is a consequence of the claim I motivated in the previous section — any such sharp response would constitute a complete categorisation of cases into 'applies', 'doesn't apply', or 'other'.

Fine's argument hopes to show that it's not possible for us to do both — that is, it's not possible to assign statuses to the sentences in a way that, while consistent with the reddest shade being red and the least red shade not being red, is inconsistent with any sharp set of responses to the sentences taken as a whole. The strategy of this argument is to take an arbitrary assignment of semantic statuses that meets the first condition and show that it is consistent with a 'sharp response', thereby showing that the second condition is not met by that assignment of statuses.

Let's go through an example of how this can be done to get an idea. Let's say we have three sentences, p_1, p_2 and p_3, representing a sort of 'scaled down' set of sentences saying of successively more orange shades that they're red. If we were to give an assignment of statuses to these sentences which satisfied the first condition I mentioned before, it would need to be consistent with us taking p_1 to be true, and p_3 to be false. Now supposing we were to take p_1 to be true and p_3 to be false (and supposing this was all we committed ourselves to), there seem to be two obvious sharp responses that we could consistently endorse: that p_1 is true, p_3 is false, and p_2 is true, or that p_1 is true, p_3 is false, and p_2 is false. But this by itself doesn't mean that every assignment of statuses that we could make more generally which is consistent with p_1's truth and p_3's falsity is consistent with our either taking p_2 to be true or taking it to be false. For one, it seems that if we took p_2 to not be definitely true, we couldn't at the same time take it to be true (and likewise for 'not definitely false' and 'false'). Why not? Because, Fine says, if you're committed to the truth of some sentence, you're also committed to its definite truth, its definite definite truth, and so on, and so in saying that p_2 is true and yet not definitely true, you're saying that it's both definitely true and not definitely true (2008, 114).[1] In turn, though, this means that the task of finding a sharp assignment of statuses that's consistent with the truth of p_1 and the falsity of p_3 (as well as the truth or falsity of any other sentences we want to commit to) is actually the task of finding a sharp assignment of statuses that's consistent with the truth of the *super-definitised* forms of any claims we're willing to make about sentences being true or false — that is, the forms of those claims with infinitely many 'definitely's in front of them.

[1] If you don't think that this is true, then Fine's argument as a whole will go through trivially just by pointing out that, assuming the assignment of statuses you've made so far is consistent, you could just add either the truth or falsity of p_2 to that assignment to get a sharp response.

Following Fine's argument, we can approach this new task by asking in turn of each sentence p_n (which says of some object that it's red) 'are the super-definitised forms of all the sentences I'm committed to consistent with the super-definitised form of p_n?'. If the answer is 'yes', we know that we could also consistently commit ourselves to the super-definitised form of the p_n in question. In our example with three sentences, p_1 would receive a 'yes' answer to this question since, whatever sentences we're committed to, they need to be consistent with p_1's truth, and so we know we could consistently commit ourselves to the super-definite form of p_1.

If the answer is 'no' for some p_n, we ask 'are the super-definitised forms of the sentences I'm committed to consistent with the super-definitised form of p_n's *negation*?' If the answer to *this* question is 'yes', then we know that we could consistently commit ourselves to the super-definitised form of the sentence in question's negation. The sentence p_3 in our example would represent a 'yes' answer to this question, so we know that we could consistently commit ourselves to the super-definitised form of its negation.

Finally, if the answer to both questions is 'no' for some sentence, then we know that we can consistently commit ourselves to the negation of both the super-definite form of that sentence as well as the negation of the super-definite form of its negation, since we know that we couldn't consistently commit ourselves to their non-negated forms.[2]

The process described above gives us a way to sort the sentences $p_1, \ldots, p_n$ that we're dealing with into three categories: the ones whose super-definitised forms we can commit to, the ones whose super-definitised negations we can commit to, and the

[2]Note that there's an implicit reductio argument at work here — we know that committing ourselves to the super-definite form of the relevant p_n would generate a contradiction, and so would the super-definite form of its negation, and we know that the (super-definitised forms of the) sentences we're already committed to are consistent, so we infer that the negations of both the super-definite form of p_n and the super-definite form of its negation must be consistent with the sentences we're already committed to.

ones for which we can commit to the negation of both of these super-definitised claims. These three categories can't consistently overlap, so if we can sort all of the sentences into these three categories, we've found a sharp response that we can consistently commit to.

Let's look at how this works for our three sentence example. As I mentioned, p_1 fits into the sentences whose super-definitised forms we can commit to and p_3 fits into the sentences whose super-definitised negations we can commit to. What about p_2? It depends which other sentences we're committed to the super-definitised forms of. Obviously I can't go through all of the cases individually, but let's work through three examples. If we're not committed to any claims at all involving p_2, then we could consistently commit to the super-definitised form of p_2. If we were just committed to p_2's not being *definitely* true, then we could consistently commit to the super-definitised form of its negation. And if we were committed to both its not being definitely true and its not being definitely false, then we could consistently commit to neither its super-definite form being true nor its super-definitised negation being true.

In any case, we could consistently commit to giving p_2 *some* status. This means that, as far as our three sentence case goes, if we're assigning statuses that are consistent with p_1's truth and p_3's falsity, there's always a sharp response that we could consistently assign which draws a division at the 'super-definite' level, so any such assignment of statuses will not meet the requirement that we need to assign statuses in a way that's inconsistent with sharp responses. And the difference between our three sentence case and a set of sentences saying of a more fine-grained set of shades that they're red, say, is only one of scale, so this argument applies just as much to those sentences. It's therefore not possible, Fine's argument goes, to assign statuses to these sentences which would be both consistent with there being some 'true' and 'false' cases and inconsistent with a sharp set of responses at some level.

In the following section I will offer a more detailed exposition of Fine's proof of this claim, or more specifically, his proof of the theorem which corresponds to it.

2.2 Some Set-Up

Fine's proof makes use of the 'D' operator that we've seen previously. As before, we can interpret this operator in a number of ways. As we will see, it can be replaced by any operator which is factive and closed under entailment, so this proof ought to work under many different ways of interpreting the 'D' operator (and in turn under many different ways of interpreting the notion of a 'borderline case'). To give some examples, we could interpret 'DA' as 'A is true' or ''A' has truth-value 1', or 'A is knowable'.

We first need some new definitions relating to the 'D' operator. First, '$D^\infty A$' is defined as $A \wedge DA \wedge DDA \wedge ...$, allowing us to express that a sentence is definite at all orders — we'll say here that '$D^\infty A$' is the super-definitised form of 'A'. Second, for any set Δ, $D\Delta$ is defined as the set $\{DA : A \in \Delta\}$ — that is, the set of sentences containing the definitised forms of each sentence in Δ — and $D^n\Delta$ is $\{D^n A : A \in \Delta\}$, the set of sentences containing the nth-order definitised forms of each sentence in Δ. In turn, we'll say that a set Δ is 'D-closed' when $D\Delta \subseteq \Delta$. We can then define the set Δ^D as the smallest D-closed set containing Δ (2008, 127). To put this another way, Δ^D is the smallest superset of Δ such that if A is in that superset, DA is also in that superset. In effect, then, if we have a set Δ, Δ^D contains the sentences of Δ plus all of those sentences made definite, definitely definite, and so on.

With this in place, Fine places some requirements on the consequence relation $\vdash$ that he's working with. This consequence relation allows for multiple conclusions to be a consequence of a set of premises, so $\Delta \vdash \Gamma$ if $\Delta \vdash A$ for each A in Γ. [3] Fine divides

[3]It is important to note that this is different from more common multiple-conclusion consequence relations where $\Delta \vdash \Gamma$ if $\Delta \vdash A$ for *some* A in Γ.

the rules that this consequence relation obeys into three kinds. The first set of rules are (2008, 127-8):

Identity: $A \vdash A$

Weakening: If $\Delta \vdash A$ and $\Delta \subseteq \Gamma$, then $\Gamma \vdash A$

Cut: If $\Delta \vdash \Gamma$ and $\Gamma, \Theta \vdash A$, then $\Delta, \Theta \vdash A$

The second set of rules are:

Reductio: If $\Delta, A \vdash \neg B, B$, then $\Delta \vdash \neg A$

$\wedge$-introduction: If $\Delta \vdash \Gamma$, then $\Delta \vdash \bigwedge \Gamma$, where $\bigwedge \Gamma$ is the conjunction of all of the sentences in Γ

$\wedge$-elimination: $\bigwedge \Delta \vdash A$ if $A \in \Delta$

The third set of rules, concerning 'D', are:

Factivity: $DA \vdash A$

Closure: If $\Delta \vdash A$, then $D\Delta \vdash DA$

Put together, these three sets of rules seem to provide a fairly 'minimal' consequence relation.[4] Fine therefore intends for the proof of the claim that vagueness is impossible to work regardless of which logic we adopt (2008, 112).

Now we'll introduce the notion of 'definite consequence'. This is at the centre of the proof, and ultimately what Fine proves is a fact about definite consequence. This relation, $\vdash^D$, is defined so that $\Delta^D \vdash \{A\}^D$ if and only if $\Delta \vdash^D A$.[5] One thing we can prove about $\vdash^D$ is that:

Lemma 1: $\Delta \vdash^D A$ if and only if $\Delta^D \vdash A$.

For the left-to-right direction, suppose $\Delta \vdash^D A$. Then $\Delta^D \vdash \{A\}^D$. A is a member of $\{A\}^D$, so $\Delta^D \vdash A$. For the right-to-left direction, suppose $\Delta^D \vdash A$. Then closure gives

[4]Perhaps, though, the reductio and cut rules in particular will come across as suspicious: these could well be abandoned by adherents of many-valued and tolerant logics respectively, for example. We'll come back to this.

[5]Fine states this definition as '$\Delta \vdash^D A$ if $\Delta^D \vdash \{A\}^D$' (2008, 128), but since he states that the left-to-right direction in the proof of lemma 1 (which we come to in a moment) is 'trivial' (2008, 128), he presumably has in mind that his definition should be read as containing an 'if and only if' rather than just an 'if'.

us $D(\Delta^D) \vdash DA$ and, in turn, $DD(\Delta^D) \vdash DDA$, $DDD(\Delta^D) \vdash DDDA$, and so on. But $D(\Delta^D), DD(\Delta^D), DDD(\Delta^D)$, and so on, are all subsets of Δ^D. So, by the weakening rule, $\Delta^D \vdash A$, $\Delta^D \vdash DA$, $\Delta^D \vdash DDA$, and so on. But this is just to say that $\Delta^D \vdash \{A\}^D$, and therefore, by definition, $\Delta \vdash^D A$ (2008, 128).[6]

What is the significance of definite consequence? In essence, saying that one set of sentences definitely entails another is just to say that if you definitised (at all orders) the sentences in the former set, they would entail the definitised (at all orders) forms of the sentences in the latter. The notion of definite consequence is interesting for our purposes because facts about definite consequence are facts about what follows from super-definitised sentences in particular, and because we're working from the assumption that if we're committed to a sentence, we're committed to its being super-definite. Talking about the definite consequence relation is going to help us because it's going to allow us to focus just on what happens at the 'top' level of definiteness, and not get caught up in the more complicated relationships that might hold between different lower levels.

Now, Fine demonstrates that definite consequence shares the following features with 'normal' consequence:

Lemma 2: $\vdash^D$ conforms to identity, weakening, and cut.

That is:

Identity: $A \vdash^D A$

Weakening: If $\Delta \vdash^D A$ and $\Delta \subseteq \Gamma$, then $\Gamma \vdash^D A$

Cut: If $\Delta \vdash^D \Gamma$ and $\Gamma, \Theta \vdash^D A$, then $\Delta, \Theta \vdash^D A$

The proof of each part of this makes use of lemma 1 and the identity, weakening, and cut rules for $\vdash$ respectively (2008, 128).

[6]Strangely, rather than 'gathering up' all of the sentences A, DA, DDA, and so on, into $\{A\}^D$ and applying the definition of $\vdash^D$ as I have in my rendering of the proof, Fine concludes his version of the proof by applying the $\wedge$-introduction rule to give $\Delta^D \vdash D^\infty A$, which wasn't what he needed to show.

The form of reductio that holds for $\vdash^D$ is the following:

Lemma 3: If $\Delta, A \vdash^D B, \neg B$, then $\Delta \vdash^D \neg D^\infty A$.

To show this, suppose that $\Delta, A \vdash^D B, \neg B$. Then $\Delta^D, \{A\}^D \vdash B, \neg B$ is true by lemma 1. Remembering that $D^\infty A$ is just $A \wedge DA \wedge DDA \wedge \ldots$ (that is, $\bigwedge\{A\}^D$), the $\bigwedge$-elimination rule gives us $D^\infty A \vdash \{A\}^D$, and so cut for $\vdash$ gives us $\Delta^D, D^\infty A \vdash B, \neg B$. But then reductio for $\vdash$ lets us infer $\Delta^D \vdash \neg D^\infty A$, which is equivalent to $\Delta \vdash^D \neg D^\infty A$ by lemma 1 (2008, 129).

The main 'impossibility' theorem concerns 'compatibility', and 'individual' and 'collective' responses to formulas and sets of formulas. A set Δ is 'incompatible' when there is some sentence B such that $\Delta \vdash^D B$ and $\Delta \vdash^D \neg B$, and 'compatible' otherwise. Put another way, a set of sentences is compatible when the super-definitised forms of all of its sentences do not entail a contradiction. Two sets Δ and Γ are incompatible when $\Delta, \Gamma \vdash^D B$ and $\Delta, \Gamma \vdash^D \neg B$ for some B (2008, 129).

Three more pieces of terminology. An 'individual response' to a formula B is a formula $A(B)$. For example, 'DB' is an individual response to B, saying that it is definite. A 'collective response' to a set of formulas $B_0, B_1, \ldots, B_n$ is a sequence of individual responses $A_0(B_0), A_1(B_1), \ldots, A_n(B_n)$. We can think of individual responses as things that we might commit to if we were trying to assign statuses to the sentences saying of shades that they're red: we might commit to p_1 (in which case A_1 itself would be 'empty'), $DD\neg p_3$ (in which case A_3 is '$DD\neg$'), and $\neg D^\infty p_2$ (in which case A_2 is '$\neg D^\infty$'), to give some examples. Collective responses are just different ways of gathering these individual responses together. Now a collective response to a set of formulas is 'sharp' if:

(1) $A_i \neq A_j$ for some $i, j \leq n$; and

(2) A_i is inconsistent with A_j whenever $A_i \neq A_j$ for $0 \leq i < j \leq n$.

'Sharp' collective responses under this definition give a response to each formula, exhibit a change in response over the sequence of formulas, and are made up of responses which are (individually) either identical to other responses or inconsistent with them. Fine takes it that theories of vagueness should not be compatible with sharp responses: if they take the vagueness of expressions to consist in some indeterminacy surrounding the range of their application, they should not allow for there to be a way to respond to each case in a gradually changing set of objects which cleanly divides all of the cases (2008, 129). The assumption that we're working under, that there's no correct way of dividing up all of a gradually changing set of objects into those which the relevant expression applies to, those which is does not, and any others, should likewise come out as inconsistent with any sharp collective responses about a set of such objects.

2.3 The Impossibility Theorem

We're now in a position to look at the 'impossibility' theorem (2008, 129):
Theorem 1: Take formulas $B_0, B_1, \ldots, B_{n+1}$, $n \geq 0$. Then there is no set of formulas Δ_0 which is compatible with $\{B_0, \neg B_{n+1}\}$ and yet incompatible with any sharp response to $B_0, B_1, \ldots, B_{n+1}$.

The strategy for the proof is to suppose there could be such a set Δ_0 and use it to construct a new set by gradually 'adding' certain formulas from $B_0, B_1, \ldots, B_{n+1}$, or their negations, to it, each time preserving compatibility with Δ_0 itself. We then use the set we end up with to define a sharp response which identifies a set of definite cases at all orders, a set of definite (at all orders) negative cases, and a set of cases which fall into neither of those two categories. From the way we construct the set which gives us this collective response, we can show that this set of responses is compatible with the set Δ_0, and therefore show that Δ_0 is compatible with a sharp response.

So, suppose Δ_0 is a set of formulas compatible with $\{B_0, \neg B_{n+1}\}$. Then we can define a set Δ_1 as $\Delta_0 \cup \{B_0, \neg B_{n+1}\}$. We then define sets Δ_2, Δ_3, and so on, by defining Δ_{k+1} for $k = 1, 2, \ldots, n$ as follows (2008, 129-30):[7]

Δ_{k+1} $= \Delta_k \cup \{B_k\}$ if Δ_k is compatible with $\{B_k\}$

$= \Delta_k \cup \{\neg B_k\}$ otherwise, when Δ_k is compatible with $\{\neg B_k\}$

$= \Delta_k$ otherwise.

Now, we assumed from the start that Δ_0 is compatible with $\{B_0, \neg B_{n+1}\}$. So Δ_1 is itself compatible. And by the definition of Δ_{k+1}, we only 'add' elements to the relevant Δ_{k+1} if they are compatible with Δ_k. So from the compatibility of Δ_1 we can see that Δ_{n+1} is also compatible. From this we can also infer that for no sentence B_k will $B_k, \neg B_k \in \Delta_{n+1}$, because Δ_{n+1} would be incompatible in that case. The way we have defined the sets Δ_k also lets us see that $\Delta_k \subseteq \Delta_l$ for $0 \leq k < l \leq n+1$, because for each set Δ_k, Δ_{k+1} is the union of Δ_k and another set (2008, 130).

Fine uses Δ_{n+1} to define a collective response to $B_0, B_1, \ldots, B_{n+1}$:

$A_k(p) = D^\infty(p)$ if $B_k \in \Delta_{n+1}$

$A_k(p) = D^\infty\neg(p)$ if $\neg B_k \in \Delta_{n+1}$

$A_k(p) = \neg D^\infty(p) \wedge \neg D^\infty\neg(p)$ otherwise.

This gives us a unique response to B_k for every k, since Δ_{n+1} is compatible, and for each k, either $B_k \in \Delta_{n+1}$, $\neg B_k \in \Delta_{n+1}$, or $B_k, \neg B_k \notin \Delta_{n+1}$ (2008, 130).

We know that Δ_{n+1} is compatible, so if we can show that $\Delta_{n+1} \vdash^D A_k(B_k)$ for $k = 0, 1, \ldots, n+1$, we can show that Δ_{n+1} is compatible with the set of responses just defined. This will in turn show that Δ_0 is compatible with this set of responses, since Δ_0 is a subset of Δ_{n+1}.

[7] Fine states the second condition in the definition that follows as '$= \Delta_k \cup \{\neg B_k\}$ if Δ_k is compatible with $\{\neg B_k\}$'. But the version I've stated here clarifies that this second condition only 'comes into effect' after considering the first condition. Without this clarification, given a set $\Delta_1 = \{B_0, \neg B_{n+1}\}$, Δ_2 could be *both* the set $\{B_0, B_1, \neg B_{n+1}\}$ and the set $\{B_0, \neg B_1, \neg B_{n+1}\}$. And without further clarification about how *this* could be so, this would just amount to a contradiction. I don't intend this to be a criticism of Fine; it's likely just a difference in presentation. I've also fixed an error from Fine's original presentation: originally the last condition said '$= \Delta_{k-1}$ otherwise', but Fine has confirmed in personal communication that this is a mistake.

As I mentioned before, for all k (up to $n+1$) either $B_k \in \Delta_{n+1}$, $\neg B_k \in \Delta_{n+1}$, or $B_k, \neg B_k \notin \Delta_{n+1}$, so if we can show that, whichever of these is the case for any given k, $\Delta_{n+1} \vdash^D A_k(B_k)$, then we'll have shown that $\Delta_{n+1} \vdash^D A_k(B_k)$ for all k (up to $n+1$).

So, first assume $B_k \in \Delta_{n+1}$. Then $A_k(B_k)$ is $D^\infty(B_k)$. B_k is a member of Δ_{n+1}, so $\Delta_{n+1} \vdash^D B_k$. $\{B_k\}^D \vdash \{B_k\}^D$ is just an instance of identity, so by lemma 1, $B_k \vdash^D \{B_k\}^D$. And $\{B_k\}^D \vdash^D \bigwedge(\{B_k\}^D)$ by $\bigwedge$-introduction. And $\bigwedge(\{B_k\}^D)$ is equivalent to $D^\infty B_k$. In sum, then, we have $\Delta_{n+1} \vdash^D B_k$, $B_k \vdash^D \{B_k\}^D$ and $\{B_k\}^D \vdash^D D^\infty B_k$. Given that $\Delta_{n+1} \vdash^D B_k$ and $B_k \vdash^D \{B_k\}^D$, we can infer (using the cut rule for $\vdash^D$), for each A in $\{B_k\}^D$, that $\Delta_{n+1} \vdash^D A$, which is just to say that $\Delta_{n+1} \vdash^D \{B_k\}^D$. Then, given that $\Delta_{n+1} \vdash^D \{B_k\}^D$ and $\{B_k\}^D \vdash^D D^\infty B_k$ we can infer $\Delta_{n+1} \vdash^D D^\infty B_k$ using cut again. We can use analogous reasoning to show that $\Delta_{n+1} \vdash^D D^\infty \neg B_k$ when $\neg B_k \in \Delta_{n+1}$ by replacing all instances of 'B_k' with '$\neg B_k$' (2008, 130).

The last possibility to consider is where $B_k, \neg B_k \notin \Delta_{n+1}$. We need to show that in this case $\Delta_{n+1} \vdash^D \neg D^\infty B_k \wedge \neg D^\infty \neg B_k$. So suppose $B_k, \neg B_k \notin \Delta_{n+1}$. From this we can infer that neither B_k nor $\neg B_k$ are compatible with Δ_k: if either had been, one would have been a member of Δ_{k+1} given the way Δ_{k+1} was defined. So for some formulas C and C', $\Delta_k, B_k \vdash^D C, \neg C$ and $\Delta_k, \neg B_k \vdash^D C', \neg C'$. From these we can use the reductio rule for $\vdash^D$ to infer $\Delta_k \vdash^D \neg D^\infty B_k$ and $\Delta_k \vdash^D \neg D^\infty \neg B_k$. Applying $\bigwedge$-introduction gives us $\Delta_k \vdash^D \neg D^\infty B_k \wedge \neg D^\infty \neg B_k$. As noted before, $\Delta_k \subseteq \Delta_l$ for $0 \leq k < l \leq n+1$, so by the weakening rule for $\vdash^D$ we get $\Delta_{n+1} \vdash^D \neg D^\infty B_k \wedge \neg D^\infty \neg B_k$ (2008, 130).

To complete the proof, we need to show that $A_0(B_0), A_1(B_1), \ldots, A_{n+1}(B_{n+1})$ is a sharp set of responses. The definition of Δ_1 already ensures that we have the first condition for a sharp response satisfied: B_0 and B_{n+1} receive different responses ($D^\infty B_0$ and $D^\infty \neg B_{n+1}$, respectively). And we can show that the second condition is satisfied by noting that each of the three responses, $D^\infty B_k$, $D^\infty \neg B_k$, and $\neg D^\infty B_k \wedge \neg D^\infty \neg B_k$ is inconsistent with each of the remaining two (2008, 130-1).

2.4 Cut and Reductio

One broad kind of objection that I alluded to in setting out Fine's argument in the first place is that we could just take issue with the structural rules that Fine take the consequence relation, and likewise the $\vdash^D$ relation, to obey. In particular I noted that the reductio and cut rules look especially suspect in the context of vagueness, taken here to be the rules:

Cut: If $\Delta \vdash \Gamma$ and $\Gamma, \Theta \vdash A$, then $\Delta, \Theta \vdash A$

Reductio: If $\Delta, A \vdash \neg B, B$, then $\Delta \vdash \neg A$

But we can adapt Fine's proof of theorem 1 in a way which makes no use of cut (nor the form of cut that he shows to hold for $\vdash^D$) as follows.

Since Fine's proof of lemma 3 involves a use of cut, we need to prove lemma 3 without using cut before getting to the proof of the theorem itself. To do this, we'll first prove the following:

Lemma 4: If $\Delta, \Gamma \vdash B, \neg B$, then $\Delta \vdash \neg \bigwedge \Gamma$.

So, suppose $\Delta, \Gamma \vdash B, \neg B$. Γ is a set made up of n many sentences, $A_1, A_2, \ldots, A_n$, so we know that $\Delta, A_1, A_2, \ldots, A_n \vdash B, \neg B$. When $n = 0$ (that is, when Γ is empty), we know that $\Delta \vdash B, \neg B$. Applying weakening to this gives us $\Delta, \bigwedge \Gamma \vdash B, \neg B$, and then an application of reductio gives us $\Delta \vdash \neg \bigwedge \Gamma$.

When $n > 0$, things get a little trickier. To show that $\Delta \vdash \neg \bigwedge \Gamma$ in this case, we first need to use induction to show that if $\Delta, A_1, A_2, \ldots, A_n \vdash B, \neg B$, then $\Delta, \bigwedge \Gamma \vdash A_1, \neg A_1$. So, first we'll establish a base case — that $\Delta, A_1, A_2, \ldots, A_n \vdash B, \neg B$ entails $\Delta, \bigwedge \Gamma, A_1, A_2, \ldots, A_{n-1} \vdash A_n, \neg A_n$. To do this, we start by using the reductio rule to infer from $\Delta, A_1, A_2, \ldots, A_n \vdash B, \neg B$ that $\Delta, A_1, A_2, \ldots, A_{n-1} \vdash \neg A_n$. Now we also know by $\bigwedge$-elimination that $\bigwedge \Gamma \vdash A_n$ because A_n is in Γ. And weakening this tells us that $\Delta, \bigwedge \Gamma, A_1, A_2 \ldots, A_{n-1}, \vdash A_n$. But, in turn, applying weakening to $\Delta, A_1, A_2, \ldots, A_{n-1} \vdash$

$\neg A_n$ gives us that $\Delta, \bigwedge\Gamma, A_1, A_2, \ldots, A_{n-1} \vdash \neg A_n$, and so we know that $\Delta, \bigwedge\Gamma, A_1, A_2, \ldots, A_{n-1} \vdash A_n, \neg A_n$.

Having shown this base case, we now need to show that, for any $m > 0$, if $\Delta, \bigwedge\Gamma, A_1, A_2, \ldots, A_m \vdash A_{m+1}, \neg A_{m+1}$, then $\Delta, \bigwedge\Gamma, A_1, A_2, \ldots A_{m-1} \vdash A_m, \neg A_m$. We do this in just the same way as the base case: From $\Delta, \bigwedge\Gamma, A_1, A_2, \ldots, A_m \vdash A_{m+1}, \neg A_{m+1}$ we can infer by reductio that $\Delta, \bigwedge\Gamma, A_1, A_2, \ldots, A_{m-1} \vdash \neg A_m$. And we know that $\bigwedge\Gamma \vdash A_m$ because A_m is in Γ. So again, we apply weakening to $\bigwedge\Gamma \vdash A_m$ to give $\Delta, \bigwedge\Gamma, A_1, A_2, \ldots, A_{m-1} \vdash A_m$, which in turn gives us that $\Delta, \bigwedge\Gamma, A_1, A_2, \ldots, A_{m-1} \vdash A_m, \neg A_m$, which is what we wanted.

Now since we know $\Delta, \bigwedge\Gamma, A_1, A_2, \ldots, A_{n-1} \vdash A_n, \neg A_n$ and that, for any $m > 0$, if $\Delta, \bigwedge\Gamma, A_1, A_2, \ldots, A_m \vdash A_{m+1}, \neg A_{m+1}$, then $\Delta, \bigwedge\Gamma, A_1, A_2, \ldots A_{m-1} \vdash A_m, \neg A_m$, we can infer that for all m such that $n \geq m > 0$, $\Delta, \bigwedge\Gamma, A_1, A_2, \ldots A_{m-1} \vdash A_m, \neg A_m$. But then letting $m = 1$ gives us that $\Delta, \bigwedge\Gamma \vdash A_1, \neg A_1$. From there we apply the reductio rule, giving $\Delta \vdash \neg\bigwedge\Gamma$. We've therefore gone through all of the cases, and so have shown that if $\Delta, \Gamma \vdash B, \neg B$, then $\Delta \vdash \neg\bigwedge\Gamma$.

Now let's use lemma 4 to prove lemma 3: that if $\Delta, A \vdash^D B, \neg B$, then $\Delta \vdash^D \neg D^\infty A$. Suppose $\Delta, A \vdash^D B, \neg B$. Then, by lemma 1, $\Delta^D, \{A\}^D \vdash B, \neg B$. Applying lemma 4 gives us that $\Delta^D \vdash \neg\bigwedge\{A\}^D$. But this is just to say that $\Delta^D \vdash \neg D^\infty A$. In turn, applying lemma 1 to this gives us $\Delta \vdash^D \neg D^\infty A$, which is what we wanted.

Now theorem 1, to restate, says: 'Take formulas $B_0, B_1, \ldots, B_{n+1}$, $n \geq 0$. Then there is no set of formulas Δ_0 which is compatible with $\{B_0, \neg B_{n+1}\}$ and yet incompatible with any sharp response to $B_0, B_1, \ldots, B_{n+1}$.' So, suppose Δ_0 is a set of formulas compatible with $\{B_0, \neg B_{n+1}\}$. We'll define the set Δ_{n+1} in the same way as in Fine's proof. Having defined Δ_{n+1} in this way, we also define a collective response in the same way as Fine. We then need to to show that, for each k, $\Delta_{n+1} \nvdash^D A_k(B_k)$, in order to show that $A_k(B_k)$ is compatible with Δ_{n+1} for each such k. As before, there are three cases for this.

First, assume $B_k \in \Delta_{n+1}$. Then we want to show that $\Delta_{n+1} \vdash^D D^\infty B_k$. We start by noting that $\{B_k\}^D \vdash \{B_k\}^D$ is an instance of identity. $\{B_k\}^D$ is just the set $\{B_k, DB_k, DDB_k, \dots\}$, so we can then apply $\bigwedge$-introduction, giving $\{B_k\}^D \vdash B_k \wedge DB_k \wedge DDB_k \wedge \dots$. But $B_k \wedge DB_k \wedge DDB_k \wedge \dots$ is just $D^\infty B_k$, and so $\{B_k\}^D \vdash D^\infty B_k$. From there we can apply lemma 1 to give $\{B_k\} \vdash^D D^\infty B_k$, and (since $\{B_k\}$ is a subset of Δ_{n+1}), weakening for $\vdash^D$ lets us infer from this that $\Delta_{n+1} \vdash^D D^\infty B_k$, which is what we wanted to show. We can apply analogous reasoning to the second case, where $\neg B_k \in \Delta_{n+1}$, to show that $\Delta_{n+1} \vdash^D D^\infty \neg B_k$.

The last case requires us to show that $\Delta_{n+1} \vdash^D \neg D^\infty B_k \wedge \neg D^\infty \neg B_k$ when $B_k, \neg B_k \notin \Delta_{n+1}$. So suppose $B_k, \neg B_k \notin \Delta_{n+1}$. Then we know, given the way Δ_{n+1} was constructed, that Δ_{n+1} is incompatible with B_k (and $\neg B_k$). So we also know that, for some C, $\Delta_{n+1}, B_k \vdash^D C, \neg C$. Here we can just apply lemma 3 (which we proved earlier), giving $\Delta_{n+1} \vdash^D \neg D^\infty B_k$. We can apply analogous reasoning to give $\Delta_{n+1} \vdash^D \neg D^\infty \neg B_k$, instead working from the fact that, for some C, $\Delta_{n+1}, \neg B_k \vdash^D C, \neg C$ (which we know because we constructed Δ_{n+1} to be incompatible with $\neg B_k$). Having established that both $\Delta_{n+1} \vdash^D \neg D^\infty B_k$ and $\Delta_{n+1} \vdash^D \neg D^\infty \neg B_k$, then, we just need to apply $\bigwedge$-introduction for $\vdash^D$,[8] giving $\Delta_{n+1} \vdash^D \neg D^\infty B_k \wedge \neg D^\infty \neg B_k$, which is what we wanted, and which means we've gone through all of the cases. Showing that the collective response that was defined is a 'sharp' response goes just the same as before, and so we have a cut-free proof of theorem 1.

This section has served to show that cut can be eliminated from Fine's proof of his theorem 1. But this by itself doesn't show that Fine's proof can be made to work using a sufficiently 'minimal' logic. For one, reductio still seems essential to Fine's proof. And if it isn't eliminable, Fine's 'impossibility' result can only be shown if we are already making some relatively strong assumptions about logical consequence. Where Fine

[8] We can prove that $\bigwedge$-introduction holds for $\vdash^D$ as follows. Suppose $\Delta \vdash^D \Gamma$. Then, by lemma 1, $\Delta^D \vdash \Gamma$. $\bigwedge$-introduction then gives us that $\Delta^D \vdash \bigwedge\Gamma$. Applying lemma 1 then gives $\Delta \vdash^D \bigwedge\Gamma$ as desired.

had perhaps hoped to show that, within any reasonable logical framework, vagueness can be shown to be impossible in some sense, what he may have demonstrated instead is that, in order to make sense of vagueness, we need to take reductio not to be a valid rule of inference. But this is a very different claim. Really what would be interesting would be an argument that showed this kind of impossibility result *without* relying on even as strong logical resources as reductio, since that would avoid the reply from any particular theorist that they were making unacceptable inferences by their lights. In the remainder of this chapter we'll look at some ways of doing this, and in the end we'll reach a form of argument that we can deploy which actually makes use of whatever logical resources any particular theorist of vagueness advocates in the first place.

3 Against 'Correct' Semantics of Vagueness

3.1 Similar Arguments

Fine's argument might not convince some theorists of vagueness that vagueness is 'impossible' in the sense he attempts to prove, given its reliance on reductio. Still, it seems that the basic idea of the proof is getting at something correct: if, at the level of meaning, vagueness is not consistent with a complete set of responses to the relevant set of gradually changing objects, then there seems to be something wrong with even incomplete sets of responses at that level, since they themselves are consistent with complete sets of responses. What we want to do, as theorists, is not only to *not* commit ourselves to any complete set of responses, but also *to* commit to the claim that *there are no such sets of responses*. If we can't commit to this claim, whatever our theory is seems to lapse into a kind of epistemicism — if we can't say that there are no complete correct sets of responses, what alternative could we hold consistently?

What a correct theory of the meaning of vague terms would have to amount to, it seems, would therefore be one that was *inconsistent* with complete sets of responses. Reaching a correct theory, then, looks like an impossible task, since both incomplete and complete sets of responses come out as incorrect (since both are consistent with a complete set of responses) — this is the thought that we want to preserve from Fine.

Indeed, Fine is hardly the first to offer an argument that (either explicitly or implicitly) taps into this basic idea. Sainsbury (1990, 3, 1992, 180-1), for one, discusses what he calls the 'transition question'. This line of reasoning also appears in Horgan (1994, 173-6), who calls it the 'forced march' sorites paradox.

The line of argument from Sainsbury and Horgan is as follows. Suppose there's someone who's 7ft tall. Probably it's true to say of them that they're tall, that they're definitely tall, that there's no indeterminacy surrounding the fact that they're tall. Now consider someone just slightly shorter than that person — what could truthfully said about this person? One of two claims about this person must be correct: either their status with respect to 'tall' is exactly the same, or something is different, whether this means they're not tall, or it's indeterminate that they're tall, or that there's no fact of the matter, to give some options. Presumably the former is true of this person, since they're only slightly shorter than 7ft. But what happens when we iterate this question, looking at progressively shorter people? It seems that either at some point it will be correct to say that something is different, or it will never be correct to say that something is different. But this poses a dilemma: if there's a change in the semantic status of sentences saying of some object that they're 'tall', say, on the basis of a small change, then the meaning of 'tall' draws a sharp boundary of some sort, (and by analogous further reasoning there's a complete categorisation of all the relevant objects), but if there's no change, then it would turn out that someone who was 4ft tall was tall, which contradicts the (obvious) claim that someone who is 4ft tall is not tall. What's more (though not explicitly mentioned by Horgan or Sainsbury), every object in the set of objects would as a result turn out to be at least one of tall, not tall, or something else, which would amount to a complete categorisation anyway.

The same idea that I outlined above is at play in this kind of argument. The first horn of the dilemma presented in this argument is unacceptable because complete categorisations of the relevant sets of objects are unacceptable. Meanwhile, the second horn shows that once you've got a *partial* categorisation (by saying that some particular objects are tall and that some aren't, in this case), adding in the assumption that there are no sharp boundaries (and by extension that there is no complete categorisation) creates problems, since it just ends up in contradictions.

Now, this kind of informal reasoning is fine for getting the basic idea across, and poses a serious challenge to any theorist about vagueness who takes the idea of vague terms not drawing sharp boundaries (or admitting of complete categorisations of the kind I've been talking about) seriously. Indeed, the problem as it's stated in this kind of 'forced march' argument is the one that should make any such theorist consider whether their theory really has the phenomenon right (and in particular whether their stance on the sorites paradox is correct), since the resources it relies on seem incredibly obvious and yet really seem not to fit together, in a way that's more robust than the original sorites paradox. However, while the dilemma posed within Horgan and Sainsbury's arguments seems to be a serious one, for this argument to be successful against any proposed correct logic or semantics for vague terms, it would need to be stated in that logic or semantics' own terms. But once we get as far as trying to state this argument within the language of certain theories, we will often find that those theories have their own strategies which attempt to 'defuse' it built into them,[9] and

[9]We can briefly sketch here some examples of this problem occurring. For one, supervaluationists such as Keefe make a distinction between the truth of the claim that making some tall people marginally shorter can make them not tall, and the truth of any particular claim that making some particular tall person marginally shorter can make them not tall (2000, 164-5), and so there's no person about whom a supervaluationist will say 'they're tall, but making them shorter would make them not tall', and so in turn no clear 'gotcha' case for Horgan and Sainsbury. Another example we could point to is Zardini's

so if it's going to be possible to show that the argument works generally, it's going to need to make specific use of the languages of any theory that it could target and show why these defusing strategies aren't good enough. Rather than try to show how Horgan and Sainsbury's arguments could be made to work in this way, we'll now go on to see a related kind of argument that can be deployed more systematically as a challenge to particular attempts to give a semantics for vague terms.

3.2 Complete Categorisations

Let's say we have a language which contains the usual connectives of a quantified logic ($\wedge, \vee, \rightarrow, \neg, \exists, \forall$), a predicate F, and an infinite number of names and variables. And let's say we add to this language an operator 'O' which operates on sentences — so if 'A' is a sentence, 'OA' is a sentence too. 'O' is going to represent the expression 'other' in this language.

Now let's say we put this language to work in trying to formulate a sketch of an accurate semantics for the expression 'tall'. Our predicate 'F' will represent the expression 'tall', and we'll take names 'a_1' to 'a_n' (whatever n might be in this case) to pick out successive objects in one of our gradually changing sets of objects, say with 'a_1' picking out someone who is 7ft tall, and 'a_n' picking out someone who's 4ft tall.

Let's suppose that 'O' really represents any and all of the 'other' semantic statuses (besides being straightforwardly true and straightforwardly false, for example)[10] that a sentence could have in the semantics that we're considering. What should this semantics say about the claim that, for any sentence 'A', either '$A \wedge \neg OA$', 'OA', or

tolerant logic (in his 2008): in this logic, for each case we can point to (starting from someone who's tall) it's true to say 'if this person's tall, then this marginally shorter person is tall too', but within this logic it's also not possible to infer from this that a clearly short person is tall because the consequence relation is not transitive — chaining good inferences of this sort together doesn't automatically result in a good inference.

[10] I also have in mind here 'being true only' and 'being false only' respectively — this will become relevant in a later footnote in relation to theories which allow for true contradictions.

'$\neg A \wedge \neg O \neg A$' is true? That is, should the semantics say that, for any sentence A, either it's true and has no other status, or is false and has no other status, or has some other status? One answer could be that yes, one of these must be true of any sentence in the language. But, given our assumption that the correct semantics for vague terms does not allow us to completely categorise the relevant sets of cases into those to which the relevant expression applies, those to which it does not, and any others, this is not the answer the correct semantics would give, since this assumption in effect says that one complete categorisation is correct, whatever it may be.

But what is the alternative? That some sentences in the language aren't assigned any semantic status within the semantics we're looking at. This is not merely to say that there are some sentences about which the (correct) semantics says 'there's no fact of the matter about the semantic status of this sentence' or 'there's nothing to be said about this sentence'; instead, it is to say that the semantics does not say *anything* about them. But this is a strange thing to appear in the correct semantics for the term we're looking at: why would it be that it can say nothing about such sentences, and yet *not* be able to explicitly say that there's nothing to be said about them?

Let's call a semantics in which some sentences are assigned no semantic status (not even 'has no semantic status') semantically 'quietist'.[11] The first thing to say about such a semantics is that it does not (in virtue of being quietist) necessarily get any of the judgements it makes *wrong*: it might get everything *that it says* correct. But here's the point: at some point such a theory will stop making claims about the gradually changing objects, but will not stop making them in a way which reflects any actual feature of the vague term in question — that is, it won't stop making claims because there's a class of sentences which can be definitively categorised as those about which there is nothing to say. It's exactly because such a semantics couldn't have an extra

[11]Ricardo Mena, in his (Forthcoming), uses this term in a similar, but less strict, way, and actually defends semantic quietism as a position about vagueness. We'll see more about this in chapter 6.

expression in its language representing the cases about which there is nothing to say that this must not reflect any actual feature of the relevant term — adding such an expression would amount to carving out a complete categorisation of cases, since all the cases it applied to could be encompassed by the 'O' operator when interpreted relative to that language. Such a semantics therefore could not be correct in the sense of both only saying correct things about the semantic statuses of the relevant sentences *and* saying everything correct that there is to be said (relevantly) about them.

On the assumption, then, that a purportedly 'correct' semantics for a vague expression is stated in a language which contains an equivalent for all of the features of the language I've defined, the correctness of that semantics can be challenged from two perspectives. Either the semantics says that the relevant gradually changing set of objects can be sorted into those to which the expression applies, those to which it does not, and any others — violating our initial assumption that no such categorisation is possible — or it does not say this, in which case it cannot be *the* correct semantics for the relevant term, since it will ultimately stop making claims about cases in an unprincipled way.

The significance of this dilemma hinges on quite a big 'if': it is only a dilemma for a purportedly 'correct' semantics for vague terms if that semantics is stated in a language which contains equivalents for all of the components of the language I defined — namely, the connectives of quantified logic, a predicate F, names and variables, and an 'O' operator — or, at least, if such equivalents could be defined in terms of other parts of that language. This condition is obviously not trivially satisfied by any possible semantics for vagueness, but we can make two promising remarks (at least, for my purposes) on this to begin with.

First, leaving the 'O' operator aside, it *is* trivial to find equivalents, at least, for predicates, the usual connectives of quantified logic, names, and variables, in most of the going semantics for vague terms. What's more, we don't even necessarily need predicates, quantifiers, names, or variables, since we could just have numbered propositions representing claims about successive objects in gradually changing sets of them, as we do in Fine's proof above. Really, the crucial (and more difficult) equivalent that we need to find (or define) is 'O'.

The second point to make is that any correct semantics *should* allow for an 'O' operator to be defined, just by 'collecting together' all of the values that aren't straightforwardly the best and worst ones (that is, true (only), false (only), true to degrees 1 and 0, and so on). If this isn't possible on some particular semantics, where are the values that are missing from its language? Still, we need to say more than this to show how such an 'O' operator is to be determined for any particular language that we might consider. In order to answer this question, let's take a detour into one particular attempt to give a semantics for vague terms, and see how this argument is supposed to apply to it.

3.3 Horgan's Transvaluationism

Given what I said in the first section of this chapter about what a 'correct' semantics for vague terms would look like, Horgan's take on the semantics of vagueness represents an attempt to give such a semantics in that it attempts to make sense of the claim that vagueness involves no complete categorisations at the level of use, and Horgan explicitly says that the correct semantics for vague terms must reflect this (2010, 76-7).

How does Horgan's theory work? He favours a logic with three truth-values: true, false, and neither. As well as a 'narrow' negation operator ('$\neg$'), this logic contains a contrasting 'broad' negation operator, '$\sim$':[12] '$\neg A$' is true if and only if 'A' is false, and '$\sim A$' is true if and only if 'A' is not true. That is, '$\sim A$' is true if and only if 'A' is false or neither true nor false. In Horgan's logic, the '$\sim$' operator cannot be 'driven inwards'

[12]Horgan actually uses these symbols the other way around, but for the sake of consistency with earlier writing, I will take '$\sim$' to be the special kind of negation.

(2010, 74). So, for example, $\sim \exists x(Fx)$ will not entail $\forall x \sim Fx$, and $\sim (Fa_n \vee Fa_m)$ will not entail $\sim Fa_n \wedge \sim Fa_m$.

Using these resources, Horgan is able to specify which principles he takes to be satisfied by vague expressions. Crucially, he takes $\exists x_n(Fx_n \wedge \neg Fx_{n+1})$ to be neither true nor false when F is vague, and so endorses $\sim \exists x_n(Fx_n \wedge \neg Fx_{n+1})$ (2010, 74). This is significant, for one, because it means that Horgan's theory is not consistent with a categorisation of all of the objects into those of which it's true to say that they're, say, tall, and those of which it's false to say they're tall, since that would make $\exists x_n(Fx_n \wedge \neg Fx_{n+1})$ true. So it's good that this claim isn't true on Horgan's theory. But what's the significance of its not being false? The fact that it isn't false on this theory means that it can block *another* way in which we might try to arrive at a complete categorisation of all the cases. Since $\exists x_n(Fx_n \wedge \neg Fx_{n+1})$ isn't taken to be false, its (narrow) negation, $\neg \exists x_n(Fx_n \wedge \neg Fx_{n+1})$, and likewise its equivalent, $\forall x_n(Fx_n \rightarrow Fx_{n+1})$, isn't true, and so we can't just conclude from the claim that someone who is 7ft is tall, for example, that everyone (including someone who is 4ft tall) is tall, thereby completely categorising all possible cases.

Now if this were the end of the story, we could apply the argument I gave above to Horgan's theory easily: we'd say that '$\sim A \wedge \sim \neg A$' was the equivalent of 'OA' in his theory — this just represents a sentence's being neither true nor false. In this case, we could then ask of Horgan's theory whether '$A \wedge \neg(\sim A \wedge \sim \neg A)$', '$\neg A \wedge \neg(\sim \neg A \wedge \sim \neg\neg A)$', or '$\sim A \wedge \sim \neg A$' is true for all sentences 'A'. If the answer to this was 'yes', then the theory would actually contain a complete categorisation of the relevant set of gradually changing objects. If 'no', it would be semantically quietist, since there would be some sentences that were unaccounted for by the theory.

Horgan is, however, aware that his theory should not divide all of the relevant cases into those to which the expression in question applies, those to which it does not, and those to which it neither applies nor fails to apply. Instead, he says that his suggested three-valued logic must be iterated: not only should sentences in the object language (that is, sentences about whether people are tall or not) be assigned one of three values, but the sentences in the language which *assigns* these values must themselves be evaluated using three values. So, for example, in this 'metalanguage', ''a person who is 5ft 10 is tall' is false' will be given a value, say true. This has the effect that sentences can be neither true nor false, neither neither true nor false nor not neither true nor false (breathe!), and so on. Horgan says that we should iterate this idea, too, to avoid iterations of the same kind of problem, creating an infinite number of metalanguages, each containing sentences which make claims about the sentences in the language 'below' it. The sentences in each such metalanguage are assigned values within the same three-valued framework. (2010, 73-4).

This detail of Horgan's account makes things more complicated, but it isn't good enough to resist the kind of argument that I'm putting forward. Iterating the three-valued framework doesn't change the fact that we can still identify an 'O' equivalent in any such iteration. In the first language we introduced, the equivalent for 'OA' is '$\sim A \wedge \sim \neg A$'. In the second language we introduced, the sentences that fall into the 'other' category also include those which not only don't fall into the 'true' or 'false' categories, but also don't fall into the 'neither true nor false' category — that is, the sentences 'A' such that $(\sim (\sim A \wedge \sim \neg A) \wedge \sim \neg(\sim A \wedge \sim \neg A))$.[13] In total, then, in this second language, sentences falling into the 'other' category are sentences 'A' such that $(\sim A \wedge \sim \neg A) \vee (\sim (\sim A \wedge \sim \neg A) \wedge \sim \neg(\sim A \wedge \sim \neg A))$.

[13] Why $(\sim (\sim A \wedge \sim \neg A) \wedge \sim \neg(\sim A \wedge \sim \neg A))$? The sentences which satisfy this expression are, on the one hand, the ones for which $\sim (\sim A \wedge \sim \neg A)$ is not true — and so are ones which can't be said to be neither true nor false. On the other hand, they are ones for which $\neg(\sim A \wedge \sim \neg A)$ is not true, and so it's not true to say that they *are* either true or false, since $\neg(\sim A \wedge \sim \neg A)$ is just the same as $A \vee \neg A$. We can see that this last equivalence holds as follows. $\neg(\sim A \wedge \sim \neg A)$ is equivalent to $\neg \sim A \vee \neg \sim \neg A$. The left-hand side of this disjunction says that it's false to say that A is *not true* — this is just to say that A is true, since 'A is not true' is by definition false just when A is true. The right-hand side of the disjunction says that it's false that 'A is false' isn't true, and by the same reasoning this means that 'A is false' must be true, and so A must be false. Put together, then, this disjunction says that A is either true or false.

We can keep going. In the third language up, sentences falling into the 'other' category also include the ones which don't fall into the set of cases just added to the 'other' category for the second language — that is, those sentences 'A' such that $\sim(\sim(\sim A \wedge \sim\neg A) \wedge \sim\neg(\sim A \wedge \sim\neg A)) \wedge \sim\neg(\sim(\sim A \wedge \sim\neg A) \wedge \sim\neg(\sim A \wedge \sim\neg A))$.

By applying the above reasoning successively, we have a procedure for working out what the 'O' equivalent is for any given language in Horgan's hierarchy of metalanguages. In each such metalanguage, there's a kind of semantic quietism at work — to avoid saying that the relevant 'O' expression in each such language exhausts all of the possibilities besides 'true' and 'false', a new language is introduced which classifies sentences in a way which goes beyond the possibilities captured by that 'O'. Now, the hierarchy of metalanguages that gets generated in this way doesn't 'stop' anywhere, but we can see that any one of them contains a definable expression corresponding to 'O', and so all of them are potentially liable to the argument strategy that I'm trying to apply, and none of them has a satisfying answer to it.

It might be replied at this point that none of the languages in the hierarchy of metalanguages, or the assignments of values that are given to its sentences, is supposed to be the correct semantics for the vague term, but rather all of them taken together are. But this doesn't really get round the problem: taken all together, there's still an expression we could define which would represent all of the 'other' values contained in each metalanguage — it would just be an infinitely long disjunction of all the 'O' expressions in each language in the hierarchy. Would every sentence be either true, false, or other, under this definition of 'other'? It's just the same as before: if 'yes', the semantics gives a complete categorisation of cases; if 'no', it's semantically quietist.

What have we just shown? That we can define a suitable 'O' operator in the language of Horgan's theory which represents the 'other' values that a sentence could be assigned, and that, as a consequence, his theory can be construed as either actually giving a complete categorisation of cases or being semantically quietist, and so in either case not providing a correct semantics for vague terms.

3.4 A General Strategy for Finding 'Other' Statuses

Let's now return to the question I posed before looking at Horgan's theory of vagueness in particular: when looking at some particular semantics of a vague term, how do we find out which expression in the language that that semantics is stated in corresponds to 'O' in the argument form I've given? While the ins and outs of this will vary from theory to theory, I suggest that we can draw out the following procedure for finding the relevant 'O' for any given purportedly correct semantics for vague terms.

We start by looking at the language in which the semantics we're looking at is stated. We need to identify what we can call, for simplicity's sake, an expression 'B' that ascribes 'borderline' status to sentences. We'll do this by identifying the 'top' and 'bottom' values of the semantics in question and defining an expression (if there isn't already one in the language) which expresses that a sentence has neither of these values. In the case of Horgan's theory, we expressed that a sentence 'A' had borderline status with the expression '$\sim A \wedge \sim \neg A$', for example, since this excludes the possibilities that 'A' could be true or false, the top and bottom (or 'best' and 'worst', if you prefer) values in this theory.[14] In this case, the 'borderline' expression could be con-

[14]Here are some more examples. In supervaluationist theories such as Fine (1975) and Keefe (2000), 'BA' will say that A is neither supertrue nor superfalse. In 'gap' theories such as Tye (1990), 'BA' will say that A has the 'gap' value I. In subvaluationist theories such as Varzi (1995, 1997), Hyde (1997) and Hyde and Colyvan (2008), 'BA' will say that A is both true and false (or perhaps more specifically, that A is true on some interpretations and false on others). In other paraconsistent approaches such as Priest's (2010), 'BA' will likewise say that A is both true and false. In any of the above cases involving true contradictions, 'BA' can count as an 'other' value (that is, we could define 'OA' as 'BA' here) even though A will always be true when BA is. This fits with my definition of a complete categorisation view as one in which, for all A, $(A \wedge \neg OA) \vee OA \vee (\neg A \wedge \neg O \neg A)$ is true, since, when A is both true and false in this case, $A \wedge \neg OA$ will be false, as OA will be true and not false. In Zardini's tolerant logic (see Zardini (2008, Forthcoming)), 'BA' will say that A has a tolerated, but not designated, value. In two-valued logics, 'B' (and so too 'O') will come out as empty — the question of whether '$A \wedge \neg OA$', 'OA', or '$\neg A \wedge \neg O \neg A$' is true of all the relevant sentences is then just whether '$A \vee \neg A$' is true of all of those sentences. In theories endorsing classical logic, for example Williamson (1994) and Sorensen (2001),

structed out of elements of the object language (since the '~' operator is used to express that a sentence is not true), but we might sometimes have to look to a language which makes claims about sentences in the object language. For example, degree theories of vagueness often have as their top and bottom values 'true to degree 1' and 'true to degree 0'. Whether this can be expressed in the relevant object language is going to be left up to the defender of such a theory, but it's clear that the 'borderline' expression for such theories should express that a given sentence has a value less than 1 but greater than 0.

Having done this, we then see whether the language that contains the expressions ascribing top, bottom, and borderline values, is itself taken to be vague. If it isn't vague, we can pose our dilemma at this point: the 'O' operator we're interested in will turn out just to be the expression denoting borderline status by itself, and we can ask of the theory whether all sentences have either a top, bottom, or other value, noting that any answer to this looks unacceptable if the theory is to be correct. If the language containing these expressions *is* vague, we then look to the next language in the metalinguistic hierarchy.[15] In this language we can define a new 'O' operator, making use of the 'B' operator we defined before — 'OA' will now be equivalent to '$BA \vee BBA$'. We can keep going in this pattern, expanding out the meaning of 'O' for each new (vague) metalanguage by adding in an extra iteration of 'B' to the disjunc-

the answer to this last question is straightforwardly yes. Likewise, in contextualist theories making use of classical logic, such as Graff's (2000), the answer to this question is yes (although how sentences are categorised varies with context). On the other hand, intuitionist logics, deployed in the context of vagueness by, for example, Wright (2001) and Putnam (1983), are also two-valued, but do not take '$A \vee \neg A$' to be valid — indeed, Putnam's solution to the sorites paradox rests on '$\exists x_n(Fx_n \wedge \neg Fx_{n+1})$' not being accepted (1983, 312-3). In, for example, Edgington's degree theory (see Edgington (1997)), 'BA' will say that A has a value below 1 but greater than 0. Smith's plurivaluationist degree theory (2008) can't exactly be said to have a 'BA' equivalent — perhaps the closest we could get would be to say that A has a 'truth profile' which contains some degrees of truth higher than 0 and less than 1 (although Smith gives reasons for thinking that this would mischaracterise his view (2008, 105)) — but plurivaluationism comes with a rejection of a 'correct' assignment of semantic statuses to sentences anyway, and I'll say more about this in the next chapter.

[15] As I noted earlier, Horgan's view would require us to do this. The strategy of treating successive metalanguages also has precedent in, for example, Keefe (2000) and Tye (1994).

tion we have already (so next we'd have '$BA \vee BBA \vee BBBA$', for example) — though the expression 'O' will very soon become impossible to write out 'fully' in any intelligible way! As with Horgan's view, we can ask of the assignments made within any particular metalanguage in the hierarchy whether they are complete categorisations or quietist, and neither answer is going to be acceptable by our lights. We can likewise pose the dilemma of whether an infinite hierarchy of metalanguages, *taken as a whole*, produces a complete categorisation or whether it is quietist: we know we can express this dilemma by defining an 'O' operator such that 'OA' is an infinitely long sentence of the form '$BA \vee BBA \vee BBBA \vee \ldots$', since we know that a 'B' operator is definable in any language in the hierarchy.

Where does this leave us? The basic form of the argument that I've outlined is this. Given the characterisation of vague terms that I gave in the previous chapter, we should think that a correct semantics for vague terms would not completely categorise objects into the ones to which they apply, the ones to which they don't, and any others. But if you have a language which contains some basic logical connectives, the resources to say of a set of gradually changing objects that each one satisfies the relevant predicate, and an expression representing that certain sentences have some 'other' status besides the 'top' and 'bottom' ones, you must either say that all of the objects are covered by the 'top', 'bottom', and 'other' statuses — in which case you've failed to produce a correct semantics, given our earlier assumption — or some objects must be missed out by those three statuses, and so your theory must be (at least implicitly) quietist, and so likewise can't be giving the correct semantics for vague terms. I've offered a way to make this argument apply generally by showing how the 'other' component of a language could be defined within different theories, and so have shown that, given the assumption that the correct semantics for vague terms would not draw complete categorisations, there could be no correct semantics for vague terms.

4 Conclusion

The considerations in this chapter have had an apparently quite ambitious target, in trying to show that there could be no correct semantics for vague terms. Having seen that, while it's possible to eliminate the cut rule from it, Fine's argument for this claim does seem to rely on the reductio rule, and having noted that Horgan and Sainsbury's arguments (which turn on similar ideas) seem not to be easily applied to all semantics for vagueness, I've shown a general strategy for deploying an argument — based on similar kinds of consideration — to particular proposed 'correct' semantics for vague terms, which makes use of the terms actually employed by any such theory, and which doesn't rely on any special logical apparatus. In the next chapter we'll go on to see how far-reaching this conclusion is, and the methodological lessons we can learn from it when theorising about vagueness.

Chapter 6

How Should We Theorise About Vagueness?

This chapter will spell out some of the implications of the results established in the previous chapter in particular. To be more precise, we're going to focus on the following questions: How revisionary is the claim that there could be no correct semantics for vagueness? How might we investigate the semantics of vague terms more generally if there could be no correct semantics for them? Since some traditional problems of vagueness are typically addressed by appealing to features of the logic or semantics of vague terms, how do we go about addressing such problems? In answering this last question, we'll address higher-order vagueness, and show just how we can avoid the paradoxes tied up with it.

1 'Incorrect' Semantics?

Let's suppose that we accept that there could be no correct semantics for vague terms for the reasons I gave in the previous chapter. Any theorist about vagueness taking themselves to be aiming to properly capture the phenomenon of vagueness might see this as a challenge to their theory as a whole (whatever it might be). By saying that there are no correct semantics for vague terms, am I saying that theorists of vagueness who are trying to give such a correct semantics need to go back to the drawing board (or give up altogether)? Not necessarily. In making this claim, I intend to make more of a methodological point rather than to criticise any particular set of theories, or to say that we should stop theorising about vagueness altogether. Let's take a look at whether it's possible to maintain this distance from ruling out existing theories altogether.

1.1 Impossibility Objection

We can start with the most extreme end of the possible revisionary implications of the claim that I'm putting forward. If there are no correct semantics for vague terms, does that mean that there *could be* no vague terms?[1] If we think of vagueness as a semantic phenomenon, and we also take there to be no correct semantics for vague terms, how can we make sense of there being vagueness at all?

Well, it's straightforward enough to show at least one sense in which vagueness and vague terms are not impossible (or don't in themselves embody any contradictions), even if there is no correct semantics for vague terms. Think again about the claim I made in chapter 4 in characterising vagueness: that when an expression is

[1]Horgan (in his 2010) reconstructs a similar objection to his view (defended in, for example, his 1995, 1998, 2000 and 2006) along these lines from Williamson (citing his 2002). Horgan's view involves the claim that vague terms are in some way inherently *inconsistent*, and he takes the objection posed by Williamson to be that, since there can be no true contradictions, vagueness must be *impossible* on his view (2010, 75). Now what Williamson actually says doesn't quite amount to this — the closest he comes to this claim is in saying that Horgan's view might appear to be that 'vague discourse *satisfies* semantic standards that are mutually inconsistent' (my emphasis), but that '[s]ince this claim involves an immediate inconsistency, it is presumably not what Horgan intends' (2002, 274), and in going on to suggest that there is some ambiguity about whether Horgan really is (or should be) claiming that vague discourse satisfies those standards. The impossibility objection can be read from this by considering why Horgan would not intend to say that vague discourse satisfies mutually inconsistent standards if that involved an immediate inconsistency: presumably the reason Williamson has in mind is that contradictions cannot possibly be true, and so for there to be vagueness something impossible would need to be the case.

vague, competent users of that expression won't, consistently with each other (and themselves), judge a relevant set of cases in a way which divides them up into the ones to which it applies, those to which it doesn't, and any others. Whether or not there could be a correct semantics for the expressions covered by this condition, it seems clear enough that some expressions could meet this condition. It therefore seems that there could still be vague expressions, even if there were no correct semantics for such expressions. What this highlights is just that vagueness should not be construed fundamentally as a feature of the 'meanings' of certain expressions; instead, the vagueness of expressions is, at root, a feature of the way they are 'used'. Of course, on this view, an expression's being vague has certain implications for what it means, and one key implication is that there is no correct semantics for that expression.

A potential objector could reply to this by saying that, in construing vagueness as a feature of the way expressions are used, rather than their meanings, I have only pushed the 'impossibility' objection back by a step, as I argued that when the use of those expressions have that feature, a correct semantics for those expressions would not divide all cases into those with top, bottom, and other, values (a claim which I made use of in arguing that there could be no correct semantics for vague terms). As such, they could argue that my characterisation of vague terms in terms of use still renders vagueness impossible, because the relevant feature of their use *entails* certain (impossible) features of its meaning. But this isn't quite right: my argument is better construed as relying on the claim that, since competent use doesn't stably divide up cases into 'applies', 'doesn't apply' and 'other' when there's vagueness, if there *were* a correct semantics for vague terms, it would likewise not divide up cases like that. And that's consistent with there being no correct semantics for vague terms at all.

In a different vein, they could also reply that, if I'm right in saying that the characterisation of vagueness that I offered in chapter 4 means that there's no correct semantics for vague terms, then that's just a reason to reject that characterisation: if you're convinced that there must be a correct semantics for vague terms, in offering my characterisation I'd essentially be trying to convince you that vagueness was impossible. This wouldn't get things right, either. It's possible to accept my characterisation and yet maintain that the semantics for vague terms do in fact divide all cases into 'applies', 'doesn't apply', and 'other'. The point that I made in chapter 5 is just that if the use and meaning of vague terms don't 'match up' in this way, it's mysterious how this happens. The question we're left with is this: is it more plausible that the meanings of vague terms completely categorise all cases, despite use failing to outwardly reflect any such categorisation, or that there is no correct semantics for vague terms after all? Much of this chapter should lend greater plausibility to the latter option.

Before coming to the next 'revisionary' problem, it's important to clarify that the view I'm putting forward is not that there is nothing correct to be said about the semantics of vague terms. Some obvious claims should obviously come out as true however we construe the meaning of vague terms — that 'someone who is 4ft tall isn't tall' should be deemed true (in most contexts), and that shorts are trousers should likewise be false. What's more, as we'll go on to see later, there are elements of vagueness that we can recover in the semantics that we settle on, such as giving an (at least partial) account of borderline cases. The point is just that if we tried to develop a semantics on which all features of vagueness in use are taken seriously and accounted for, we wouldn't be successful.

1.2 Are Popular Theories of Vagueness Valuable?

Does the claim that there is no correct semantics for vague terms nonetheless mean that, since many popular theories of vagueness aim to correctly capture the phenomenon of vagueness, all of these theories are worthless since they're all strictly speaking incorrect? The answer to this question must also be 'no'. As I just noted, even if there is no correct semantics for vague terms, there are still plenty of true things we can say about their meanings.

Still, there's a tension here: if these theories really are trying to get at the correct semantics for vague terms, their defenders all seem to be doing *something* wrong (notwithstanding the 'first-order' ways in which they might be getting things wrong by each others' lights). But we can construe this 'something' as big or small. We'll now look from a very general perspective at how we can do some 'damage control', and at some ways in which we could make sense of any popular theory of vagueness that takes itself to be correctly capturing the semantics of vague terms, while retaining as much as possible of the general 'shape' of those theories.[2]

2 Complete Categorisations and Semantic Quietism

In the previous chapter, I established that semantics for vague terms ultimately collapse into two different kinds. The first kind completely categorise all of the objects in a relevant set into those to which that term applies, those to which it does not, and any others. The second kind are 'quietist' in the sense that by the lights of those semantics, some objects fall outside of the categories of 'applies', 'doesn't apply, and 'other', and nothing more is said about them. While destructive to the idea of a correct semantics for vague terms, this helpfully distinguishes the two kinds of semantics we could look at once we've acknowledged that we're not aiming at a correct semantics — they're the two ways we can say at least some correct things about the terms' meanings. So, any semantics for vague terms will either find a way to categorise all of the relevant objects, or it will stay silent in some important ways about some of them.

[2]In this chapter I'm not going to make any judgement in favour or against any particular theory of vagueness on its own terms. My intention here is more to give a general outline of the options available to any theorist who wants to adapt their view to fit with the claim that there is no correct semantics for vague terms.

2.1 Complete Categorisations

Which should it be? First it might be a good idea to get a handle on what each of these kind of semantics would actually look like, and how viable either is by its own lights. Let's think first about a semantics which makes some complete categorisations of the relevant objects. The 'easiest' version of this kind of semantics is one which borrows a logical basis from epistemicism. Such a theory would say that, for every vague term, all of the objects in the relevant set of gradually changing ones are either those to which it applies or those to which it does not. This yields a very 'classical' picture, on which (as far as vagueness is concerned) the law of excluded middle holds (so, for any predicate, either it applies to any particular object or does not), and modus ponens, reductio, and double negation elimination are all valid. But this doesn't need to come along with the rest of the theory of epistemicism itself — we could just as well endorse this picture while saying that we *could* in principle know how the relevant term applies to all the relevant objects, and indeed could stipulate for ourselves exactly how they apply.

At this point it would be fair for someone not already inclined towards epistemicism to object that just because we're not aiming at a completely correct semantics or logic for vague terms, that doesn't mean that we should pretend that they're not vague at all, or that there's nothing at all to be said about how their vagueness affects their meanings. But using classical logic has at least some things to recommend it. For one, we can use this to make sense of an issue I raised early on in chapter 5, that anyone who is certain that classical logic is correct (for example epistemicists such as Williamson (see for example the stance he takes on p265-9 of his 1997) and Sorensen (see for example 2001, 1)) will steadfastly resist the move from saying that there are no stable, complete categorisations marked in *use* to saying that there are no complete categorisations at the level of *meaning* either (and so will resist the claim that there is no correct semantics or logic for vague terms), since this is just to say that classical logic must be false where vagueness is concerned. We're ultimately going to respond to this by saying that those who are set on classical logic can accept both the claim that there is no correct semantics or logic for vague terms, and yet continue to 'endorse' classical logic in some other way — we'll come on to some different ways in which we might make sense of this soon — but for now I'll just note that approaching things classically from this perspective would allow us to keep the theoretical virtues of classical logic while not requiring the additional leap of saying that vague terms actually completely categorise cases at the level of meaning.

If we're 'endorsing' a kind of classical logic, how should we say the complete categorisations that are drawn, are drawn? Where they're drawn is, to some extent, going to be arbitrary, since they're not drawn at the level of use. But at the same time our theory should say that as many clear cases of 'bald' are actually bald in our theory, and that as many clearly not bald people aren't bald, as we can: it seems that what we should be aiming at is the closest we can get to getting all of the cases correct. Eklund puts forward a suggestion along these lines, though not under the assumption that there can be no correct semantics for vague terms, roughly that the correct semantics for those terms is the one which *most closely* captures the principles of use which constitute them, since he takes it that these principles cannot all be jointly satisfied. As he puts it, 'the actual extension of a vague predicate is... what comes *closest* to satisfying the meaning-constitutive principles associated with the expression... provided any assignment of semantic values comes sufficiently close' (2005, 42-3, his emphasis).[3]

[3] We came across Eklund's view briefly in chapter 4. In talking about the principles which characterise their use he is talking primarily about the principle that speakers should be disposed to take these expressions to be tolerant (in Wright's sense, where a small change of the relevant sort can't make the difference to whether an expression applies). He takes this to mean that the principles constitutive of vague expressions cannot be jointly satisfied, since the sorites paradox exactly seems to show that if you make judgements on the basis that vague expressions are tolerant, you run into contradictions. As such, he thinks that the correct semantics for vague terms is the one which comes closest to satisfying the principle that they are tolerant. We might take issue with the claim that vague expressions are the

Why isn't this suggestion from Eklund correct? If the reasoning that I gave in the previous chapter is correct, that neither a complete categorisation of cases at the level of meaning nor a kind of quietism could really be an accurate reflection of the use made of vague terms (since either would be, at least to some degree, at odds with how the terms are actually used), why isn't it that the complete categorisation that reflects the *closest we could come* to accurately reflecting vague language use (if there is one) is the correct picture? Well, in what sense would it be the correct picture? Built into the way that it's defined is just the idea that it doesn't reflect vague language use in an entirely accurate way, so why think that it entirely accurately reflects the meaning of vague language? The closest we could come to accurately reflecting vague language use is just that — the closest thing. It's an extra leap to infer that the closest thing must be correct; what could we say to justify that inference? Eklund says that this question is not just one that needs to be answered when we're looking at vagueness: if we're trying to avoid inconsistency in semantic theories more generally and yet find that competent speakers make judgements in mutually inconsistent ways, as theorists we need to develop a semantics which *best fits* the speaker behaviours we've observed, and this is just common practice (2005, 51). This is well and good, but even with this methodological consideration (that theories ought to aim for 'best fit' when confronted with seemingly inconsistent linguistic behaviours) in mind, it's not clear why theorising in this way would deliver *correctness* (rather than just the best theory), and we'll see soon that Eklund's proposal can be re-interpreted in other terms than as getting the correct semantics, and that this re-interpretation can better 'make sense' of this tension.

ones which we are disposed to think are tolerant, but the same kind of reasoning might be thought to apply when thinking about how no complete categorisations of cases could be correct — in this case the correct semantics would be the assignment of semantic statuses that came closest to speakers' judgements overall.

Coming back to the basic 'classical' picture outlined above, we could also adopt more complicated theories, still under the assumption that there are some complete categorisations of cases. For example, thinking back to Horgan's three-valued logic, we might endorse this logic in its non-iterated form, saying that all of the cases can be divided either into those of whom it's true to say that they're bald, say, those of whom it's false to say they're bald, and those for whom it's neither true nor false. We could also endorse an iterated form of this logic, in which sentences saying of sentences that they're either true, false, or neither true nor false, are themselves assigned one of these three truth-values (and sentences saying of *those* sentences that they're true, false, or neither are themselves assigned either true, false, or neither, and so on). As we saw in the previous chapter, though, in order to prevent this view from lapsing into quietism, one such iteration of this logic would need to assign values in a precise way. Still, doing things this way has some advantages over the classical way of doing things described above, since it could perhaps better capture the idea of objects or sentences having borderline status — on this view, there are some actual gaps between full truth and full falsity when it comes to vagueness, and we might think that anything short of this steamrolls over the phenomenon altogether for the sake of a simple logic, absent the epistemic component, say, of epistemicism. It also arguably deals better with the sorites paradox, since it would allow us to say that it's not true that there's a hair that you can add to someone's head that stops them from being bald, for example — I outlined the strategy for doing this that Horgan adopts in chapter 5.

On the other hand, Eklund points out that these theoretical gains are a little superficial. In this particular case, we can illustrate the point by noting that, even if Horgan's view can retain some notion of borderline status, since we can't have an infinitely iterated form of Horgan's logic in which each iteration assigns truth-values in a *vague* way (without being semantically quietist), we can't fully capture the idea that, when we've got a vague term, all of the borderline categories (at any order) associated with that term should not be sharply bounded. And likewise, even if a three-valued logic would let us resolve the sorites paradox more intuitively, it would not allow us to solve every iteration of the sorites paradox in the same way — again, since the logic becomes precise at some point in its iteration, at some point we are going to have to say that someone could go from being bald and definitely bald at all orders to being borderline bald at *some* order, just by the addition of one hair. Eklund's point is that, if we're going to have complete categorisations in our semantics, we're going to have to say *something* counterintuitive when assigning values to sentences, and that's why he favours the simplest (classical) approach to doing this (2005, 43-4,53-4).

I'm not going to comment on whether Eklund is right about this, but whatever we think about the advantages and disadvantages of approaching complete categorisations classically or non-classically, it's worth noting at this point that there are plenty of ways to fill in a complete categorisation. Most (if not all) contemporary theories of vagueness can be readily adapted by making sure that, whatever their 'other' value is, all the cases are either covered by it, or the 'top' or 'bottom' values assigned by that theory (for example 'true' and 'false', or 'true to degree 1' and 'true to degree 0').

Before we move on to semantic quietism, let's think about a contemporary example of a 'complete categorisation' theory. Zardini's theory of vagueness deploys a 'tolerant' logic in order to try to solve the sorites paradox while maintaining that vague terms are not sensitive to small changes. In tolerant logics, the consequence relation is non-transitive: put simply, A can entail B and B entail C and yet A fail to entail C. In Zardini's case, this is implemented by distinguishing between two kinds of values that a sentence can have: 'designated' and 'tolerated'. As he puts it, these can be thought of as the 'very good' values and the 'good enough' values. All very good values are good enough, but not all good enough values are very good. The idea of his implementation of tolerant logic is that if some sentence A has a very good value, and A entails B, then you can only infer from this that B has a *good enough* value (while it may in fact have a very good value as well). In turn, though, a sentence B having a merely *good enough* value doesn't allow you to infer that C has a good enough value as well, even when B entails C (2008, 347-8). Zardini is thereby able to say that, for example, 'if you add a tiny quantity of white paint to a tin containing red paint, you will still have red paint' has a designated (that is, very good) value and yet resist a sorites argument moving from this, coupled with the assumption that some particular tin contains red paint, to the conclusion that a tin with pink paint in it actually contains red, since the step-by-step inferences we'd need to make to reach this conclusion will not all be guaranteed to start from very good values — some will only be good enough.

The sense in which this is a 'complete' categorisation view is subtle, since it doesn't seem to draw lines between 'red' and 'not red', for example, and so there's no falsifying instance of the claim that 'red' is not sensitive to small changes. But, as Zardini presents it in his (2008), at least, the important detail to notice is that, since the tolerant logic is itself stated within a *classical* logic (2008, 341), the sets of designated and tolerated values are sharply bounded: values are either in them or out of them. We therefore have a complete categorisation of cases — of the sentences saying of each tin that it contains red paint, some will have a designated value, some will have a tolerated (but not designated) value, and others will not even have a tolerated value.

2.2 Quietism

Let's now move on to think about semantic quietism about vagueness. In the last chapter I introduced quietism as being a semantic theory which, for the relevant 'O' in the language of that theory (denoting the values a sentence can have besides being 'straightforwardly' true or false), failed to make it true that, for all sentences A, $(A \wedge \neg OA) \vee OA \vee (\neg A \wedge \neg O\neg A)$. But we can also think of 'quietism' as an umbrella term for a broader range of views still. As we'll understand it here, the crucial detail of semantic quietism that distinguishes it from 'complete categorisation' views is that, while complete categorisation views take a stance on the status of all of the relevant sentences, semantically quietist views take no stance on the status of some sentences. Perhaps the simplest implementation of semantic quietism about vagueness would be to construct a theory which, for example, categorised people into two categories (which were not mutually exhaustive): ones who 'tall' applied to, and ones who it didn't — the rest it would stay silent about. Such a theory would *not* make the claim that those of whom it says 'tall' applies in the series are the only people in that series to whom 'tall' applies, although it would claim to have correctly identified any tall people that it *did* identify (and the same goes for 'not tall'). If the theory claimed to have identified all and only those people who are tall (and all and only those who are not), it would not be a semantically quietist theory in any interesting sense, since it would effectively categorise all of the people into those to whom 'tall' applies, those to whom it does not, and any others. So any quietist theory must aim not to deliver the correct verdict on at least *some* cases.[4]

Thinking about 'quietism' in the broader sense just outlined, Mena (in outlining his particular notion of quietism) suggests that there could be quietist theories which, for example, sort all cases into *three* categories: those to which the relevant vague term applies, those to which it does not, and those about which the theory explicitly goes 'silent' (Mena, Forthcoming). Now, as an implementation of quietism, it's tempting to think that this skirts too close to just being a complete categorisation theory — since all cases are accounted for by the theory — and therefore seriously blurring the lines between quietism and complete categorisations.

[4]Note though that this is slightly different in the case of terms like 'early thirties' which have a 'sharp boundary' drawn into their range of application, since (in this case) everyone younger than thirty isn't in their early thirties, and everyone who's exactly thirty is. The crucial point is that the quietist shouldn't claim to have identified all of the cases about which it's correct to make either a positive or negative claim.

To address this worry, it's helpful to think about the difference between quietist theories and complete categorisation theories not necessarily as a *formal* difference, but necessarily as an *interpretative* one. A quietist theory might properly involve silence on the part of the theory about some cases, or it might contain formal elements (such as 'values' assigned to sentences) which represent that there's officially no stance taken by the theory about those cases. On a formal level, the latter kind might look exactly the same as a complete categorisation view, but the difference will be that the complete categorisation view will assign *significance* to some sentences within that view that the corresponding quietist view will not — the quietist view will explicitly refuse to say anything significant about the semantic status of those sentences.

Perhaps a more stringent quietist would see this last kind of claim as not really being in the spirit of quietism — if a theory is saying 'there's nothing to be said' about some cases, they might say, then they're still saying *something* about them. There's also a possible kind of view that we could hold, then, call it 'strict quietism', which says that to have a quietist theory, at both the formal level, and at the level of interpretation, you must make absolutely no significant claims of any sort about the status of certain sentences. All such strict quietist theories will be quietist in the broader sense I've been describing.

One example of a 'strict' quietist theory might be a suitably adapted form of intuitionism about vagueness, as defended by, for example, Wright (2001) and Putnam (1983). Intuitionistic logics assign two truth-values to sentences — true and false — but in intuitionistic logic it isn't provable (nor is it an axiom) that all sentences are either true or false. To focus on Wright's theory of vagueness (and his deployment of intuitionistic logic), in this theory truth is tied closely to knowability: on his view, if a sentence is true, it's possible to know that it's true (2001, 59). Alongside this, Wright deploys a notion of 'borderline case' under which a borderline case of some expression is a 'quandary' — a case about which 'we do not know, do not know how we might come to know, and can produce no reason for thinking that there is any way of coming to know what to say or think, or who has the better of a difference of opinion' (2001, 71).[5] With these two aspects of Wright's view in mind, we can see that there are going to be some cases (the borderline cases) such that the intuitionistic logic being deployed will neither say it's true or false of them that the expression applies to them, nor will it say anything else about how they stand with respect to that expression (2001, 79).

Mena argues that intuitionist theories such as Wright's are not quietist[6] because they have principled reasons for failing to reach verdicts on certain sentences (namely because they're quandaries and truth requires knowability), where quietist theories don't fall silent on particular cases for principled reasons; in a quietist theory, it's *expected* that there will be some cases which could be known to satisfy the relevant expression and yet the theory still fall silent on them (Mena, Forthcoming). Mena's point seems right. To put this another way, while Wright's intuitionism is 'formally' quietist in the sense of assigning no truth-values when it comes to quandary cases, at the level of interpretation we can take its failure to do so as saying something significant about those statuses: the reason we are quiet about them is because they are quandaries, and not because we have nothing to say. This seems not to be in the spirit of quietism because such an intuitionist theory seems to say something significant about all of the sentences about whose truth-values it goes quiet.

Still, treated just as a formal framework, more stringent quietists might take intuitionistic *logic* to be more in the spirit of semantic quietism than a framework which has a separate category of sentences about which there is explicit silence, and might likewise say that it could be deployed without the particular notion of truth that's bound up with intuitionism itself. The more interesting aspect of intuitionist logic for quietists is that it represents a way to go silent about cases without attaching some

[5] See, for example, Greenough (2009) for discussion of this account.

[6] This point applies just as much whether we're interpreting 'quietist' strictly or broadly.

other value to them at the formal level: sentences can be true and false, and those exhaust the values that sentences are taken to have.

Thinking about quietist theories more generally, these theories can go beyond classifying cases as positive, negative, and 'silent' (if we think of this as a category). For example, there is room within such theories to categorise some cases as borderline cases of the relevant expression. In such theories, there'll then be either 3 or 4 categories: people who are, say, tall, people who aren't, borderline cases, and the rest (who the theory is silent about, if that counts as a category). How we want to make sense of the notion of a borderline case is up for grabs here: we could borrow resources from any theory of vagueness as we see fit. As before, it's important to note that any quietist theory which identifies some borderline cases shouldn't be aiming to get a fully complete categorisation of all borderline cases by doing this (even if it hopes to get everything it says correct) — in identifying all of the borderline cases, it would also identify all non-borderline cases, in which case such a theory would lapse into a complete categorisation theory. It's also important to point out that the category of borderline cases drawn by a theory of this kind will not overlap with the set of cases about which the theory is silent, although some cases about which the theory is silent may in fact be borderline cases (just as some may be positive cases and some may be negative cases).

To finish this introduction to quietist theories, let's think briefly about how quietists can go about developing a logic or semantics. The task of developing a logic or semantics (and in particular the task of saying how the truth-values of complex sentences are determined) is perhaps easier for the kinds of non-quietist view I sketched earlier, since on any such view a logic or semantics could be developed that makes use of the values that that view assigns to atomic sentences as a basis, where 'silence' in the questist's case seems more difficult to use as such as a basis. This is mitigated in the case of less strict forms of quietism, since they treat 'silence' as a sort of value in itself that sentences can have. In which case, there are a number of logical and semantic options available to the quietist. We could give as a basic example a quietist adaptation of a three-valued logic, where the third value is interpreted as explicit silence on the part of the theory, rather than, say, 'neither true nor false', and this interpretation applies just as much when it comes to building up complex sentences (that is, when thinking about the meanings of the connectives of the theory's language, such as 'and', 'or', 'if. . . then', and so on).[7] So, for example, in many three-valued logics, an 'and' sentence comes out as neither true nor false when at least one side of it is neither true nor false (assuming neither side is false). A quietist theory which replaced 'neither true nor false' with 'silence' would therefore be silent about the sentence 'Emeralds are green and the Information Commons is blue' (assuming the Information Commons is a case about which the theory is silent), since the theory's silence about the second half makes it impossible for the theory to reach a verdict about the sentence as a whole. This theory would likewise be silent on the status of, for example, the principle of excluded middle, because the theory wouldn't place some sentences in either the true or false categories. And this is by no means the only option for quietists — it's up to quietists to decide what their logic ought to look like.

Before moving on, let's recap the key differences between theories which offer complete categorisations of all cases for vague terms, and quietist theories. At the formal level, complete categorisation views assign a (formally) significant status to all relevant sentences, and at the level of interpretation they interpret those formally significant statuses as making a significant claim about the status of each such sentence. Meanwhile, while quietist views (broadly speaking) *can* assign significant statuses to

[7] We might think of this example as drawing from the logics developed in, for example, Horgan (2010) (which we saw some features of in chapter 5) and Tye (1994). In Tye's case, he takes the third 'value' in his theory not to be a truth-value, but rather the lack of a value (1994, 194). Semantic quietists could make similar claims to this in explaining the presence of values in their theories representing silences (if they have them) — that sometimes it's necessary (or even just more convenient) to have features of a logic which represent absences, and in the case of quietism, absences of verdict.

all of the relevant sentences at the formal level, at the level of interpretation they will say nothing significant about some of those sentences. The more strict form of quietism is just the same as this, except significant statuses will also not be assigned by such a theory at the formal level. With this in mind, if the conclusions I reached in chapter 5 are correct, then theories offering complete categorisations are bound to incorrectly specify the status of some cases, while quietist theories are bound not to *correctly* identify the statuses of some cases altogether.

3 Approaches to Logic and Semantics

Having a broad understanding of the two kinds of view we could adopt, we can now look at the reasons we might have for adopting either, and the reasons for adopting one over another, in some detail. Both of these kinds of reason turn on the reasons we have for developing logics and semantic theories in the first place. This is the topic we'll now turn to.

Things get complicated for us in thinking about our reasons for developing logics and semantic theories once we entertain the idea that our project is not merely to 'get the correct logic or semantics for natural languages like English' — which is exactly what we do have to entertain if I'm correct in saying that there's no correct logic or semantics for any vague language. Shapiro (2014) offers a helpful starting point for thinking about our goals, outlining three broad perspectives on the question of what our goals should be when developing logic and semantics (some of which can overlap).[8] I'll sketch each of these perspectives, and, in light of the above discussion of complete categorisations and quietism, I'll offer some approaches to the topic of theorising about vagueness that we could pursue from each of these perspectives. The

[8] Actually Shapiro talks about four — the fourth being the 'propositional' approach. I'm not going to directly address this approach, but I'll note here that Shapiro makes some promising comments that could be applied to the topic at hand (2014, 46).

hope is that this will leave things quite open-ended — I want to leave it open to anyone adopting any of these perspectives that they could endorse either a theory that completely categorises any relevant set of objects, or goes quiet about some cases.

The first kind of approach to consider is what Shapiro calls the 'prescriptive' approach. On this approach, theorists work from the assumption that the languages that they study are in some way not fit for purpose, and instead develop formal languages to replace them. In doing this, they will prescribe some rules for good reasoning in these new languages. Shapiro takes Tarski (in, for example, his 1956) and Frege (for example his 1879) to be defenders of this approach, as well as saying that Quine's project of 'regimenting' natural languages (for example in his 1957) falls under this heading, since a regimented language in Quine's sense will lack, for example, the vagueness and ambiguity of its corresponding natural language (2014, 45-6).

This approach takes it as a given, then, that languages such as English are deficient in some ways. Now what counts as a deficiency depends on what we're trying to achieve. Prescriptive approaches sometimes take English to be specifically not fit for describing reality, for example (see Quine (1957, 7-8)). The reasons why natural languages aren't taken to be the right ones to use are varied. Such reasons might include these languages being imprecise, ambiguous, or otherwise allowing for misunderstanding, as well as the fact that certain paradoxes and inconsistencies can apparently be generated within them. The liar sentence (a sentence saying of itself that it is false) and its variants are classic examples of such inconsistencies that can be brought to light in English. Some problems of vagueness, including the sorites paradox, can likewise be seen as fitting into this category, since they revolve around apparent contradictions in vague languages.

The claim that there is no semantics correctly reflecting vague language use can be readily accommodated within the prescriptive approach. Indeed, those following this approach might go so far as to say that the fact that we can't 'make sense' of vague terms in this way is just a reason to think that vagueness is in some way a deficiency of vague natural languages, since this leaves the 'correct' rules for reasoning impossible to work out when we're dealing with vague terms — almost all logical rules are potentially threatened by the problems of vagueness. Their task is then to prescribe some new way of reasoning with these terms.

How does this fit in with complete categorisations and semantic quietism? Either approach to giving a semantics for vagueness could be made to work within a prescriptive view. In the first instance, complete categorisations could be prescribed as the ones we ought to draw in our reasoning (whatever purpose we might have for this reasoning). This could involve a simple way of categorising cases — for example saying that, when thinking about baldness, everyone either is or is not bald, with no space in-between — or it could be more complex, involving borderline statuses, degrees of truth and falsity, and gaps between truth and falsity, for example. A number of factors could decide between doing things in any of these ways, and in large part this is going to come down to what the semantics or logic in question is being deployed for. In the context of some scientific theories, for example, we might think we only have the most accurate theory when vagueness and its related symptoms have been eliminated from our terms altogether, and this might be a reason to favour a more simple (classical) approach. On the other hand, if we just want to prescribe rules for reasoning which deviate as little as possible from actual vague language use, we might want to allow room for grey areas and borderline cases, while still officially maintaining that every sentence in the language has a defined status of some sort — and many non-classical approaches to vagueness can be modified to fit that purpose.

We could also be prescriptive quietists, maintaining that the correct rules for reasoning require us to make no judgement at times. As I mentioned before, this could be spelled out in a number of ways and could borrow resources from a number of different theories, with different attitudes towards borderline cases and the correct rules of inference — the crucial detail is just that there is sometimes nothing that can 'officially' be said. Indeed, a prescriptive quietist theory could well be developed along the same lines as the 'complete categorisation' implementation of the prescriptive approach: such a theory could carve cases up into those to which 'bald' should be said to apply, those to which it should be said not to apply, and cases about which there's nothing that should be said. Such a view would differ from a complete categorisation view just in prescribing silence for a specific set of cases rather than some set of judgements. There could also be 'strict' prescriptive quietists, who prescribe, for example, saying that some specified set of cases are tall (say), that some are not, and then make no prescription about the remaining cases.

The second approach that Shapiro describes takes formal languages to aim to 'describe' or 'represent' natural languages, and to capture their underlying forms or structures. On this kind of approach, theorists will presumably not aim to arrive at a better language than the one that they started with, but rather to arrive at an accurate representation of the target language. Shapiro takes Montague (in for example 1974), Davidson (in for example 1984), and Lycan (in for example 1984) to be defenders of this kind of approach (Shapiro 2014, 43-4).

This approach seems to be affected more acutely than the prescriptive approach when it comes to making room for there being no correct semantics for vagueness, since the big problem highlighted by this result is that there's no way of accurately representing vague natural languages formally. If the aim of this approach is to get an entirely accurate representation of these languages, then, the claim that there is no correct semantics for vagueness amounts to an argument against that approach — it means that the project can't achieve its intended goal. This is quite an interesting result in itself, as this sort of 'descriptive' or 'representational' approach to logic and semantics might have seemed like a fairly natural approach to take. Still, the claim about vagueness that I've been arguing for is surely not enough to render the broad aims of this descriptive or representational approach altogether fruitless or not worth doing, even if it does tell against the approach itself.

In starting to think about a different approach which makes sense of the broad aims of the descriptive approach, we might say that, when devising logics and semantics, our aim should be to get the *most* accurate representation of the target language possible, or the *best* description possible, whether or not an entirely accurate representation or description is possible. This way we could go about our theorising as normal, while acknowledging that vagueness is something of an anomaly that in some way resists representation or description. Of course, something more needs to be said about this — what do the most accurate representations or descriptions of a vague language look like if they don't represent vagueness entirely accurately?

The approach to formal languages that we end up with by shifting our focus to the most accurate representation or the best description is the next approach that Shapiro talks about. On this approach, theorists take themselves to be providing a 'mathematical model' of their target language, displaying or representing features of that language using mathematical tools, while not always intending to be entirely accurate representations of those languages (just as the modified descriptive approach I just outlined would not intend them to be). The models constructed on this approach will leave aside certain features of natural language and simplify or idealise others, depending on the particular purposes of those models (2014, 46-7). To get an idea of how this kind of model-building works, let's look at another project that we could describe as mathematical modelling. We might take different kinds of geometry, for example, to be mathematical models of physical space. To consider a particular case, in Euclidean geometry, lines are taken to have no width. But we can understand this as a simplification of the way physical space really is. The physical objects we represent using lines in geometry will usually have some width, but it will often be useful not to consider these widths when proving other facts about those objects using geometry, even if just because fewer steps are needed in proofs, for example.

The task for anyone taking the 'mathematical model' approach to formal languages is to find the best way to eliminate some or all features of vagueness from their models of vague natural language — if (as I've argued) a correct semantics for vague terms (that is, one which accurately reflected the use of those terms) would need to be neither quietist nor draw complete categorisations, then any model devised on this approach would need to give a semantics which fails to reflect at least some features of vague language use.[9] Both theories which draw complete categorisations and theories which are silent about some cases (whether this is interpreted 'strictly' or not) seem to be good ways of doing this, on the face of it, since neither try to capture fully the sense in which vague terms don't draw complete categorisations. We have the same advantages and disadvantages as before for adopting either of these theories — while, to give some examples, a 'classical' model which categorises all sentences into true and false would (arguably) provide a simple and intuitive picture of the 'correct' logic, a quietist model would have the advantage of not getting any of its claims wrong (assuming it's an adequate model by quietists' lights) in refusing to say anything about some cases. It might also be the case that different models are more useful in different contexts. So, for example, a classical model might be more useful for the reasons described above when we're looking at an aspect of language where we're certain that vagueness isn't relevant, and a quietist model might be more useful in representing vague language use where speakers literally go quiet about some cases.

[9]The application of this approach to vagueness (while not explicitly under the same banner) has some ancestry in Cargile's (1969). As he puts it, in describing what he calls a 'nominalistic' approach to a form of the sorites paradox, 'rather than there being some unknown instant at which [a creature] ceases to be a tadpole, it is just that, if we are making use of logic, we may be forced to choose some instant arbitrarily to be the instant when [it] ceases to be a tadpole, in much the same way that we may have to assume, in applying the differential calculus to a physical problem, that matter is infinitely divisible' (1969, 200).

A defender of the 'modelling' approach to logic and semantics might at this point face an objection. Consider again the case of lines in Euclidean geometry. There seems to be a disanalogy between this case and eliminating some essential features of vagueness in a model of a vague language. We could perhaps in principle take lines to have width in some systems of geometry without automatically generating any contradictions, but our endorsing either a quietist or 'complete categorisation' model seems to be a choice that we're forced into: the assumption that lines don't have width could perhaps be gotten rid of — and so could be legitimately used to model space more effectively — while there seems to be *no way not to* endorse quietism or a complete categorisation, both of which we know not to yield a correct semantics. So, defenders of modelling might face the objection that it is illegitimate to eliminate vagueness from their models, since the fact that they *need* to eliminate it shows that there is something systematically wrong with their model as a whole, and perhaps even that they shouldn't attempt to model vague languages in the first place. By way of reply on behalf of these theorists we can point out that their aims in modelling natural languages are to capture the greatest proportion possible of the claims that speakers commit themselves to and the inferences they make, and to describe the structure of those speakers' language use in the most accurate way possible. We can say that eliminating vagueness in some way allows these theorists to do this, since this allows them to capture a sizeable portion of the commitments and inferences that speakers typically do make, while only getting wrong (or being silent about, depending on which kind of model they're using) some cases and some inferences.

Let's look at two examples of the modelling approach in action. Edgington develops a theory of vagueness in which the values assigned to sentences come in *degrees*: sentences can can be clearly true (and so have value '1'), and clearly false (value '0'), but they can also take as values any of the real numbers between 1 and 0, signifying degrees of closeness to clear truth. In this theory, as you add more drops of white paint to the pot of red paint, at some point the truth-value of sentences saying 'this pot contains red paint' goes below 1, and gradually goes down to 0. Edgington can probably be accurately taken to be following the modelling approach in developing this theory — she says that the values that get assigned to these sentences are instrumental, and not meant to exactly represent how close they 'really' are to clear truth (1997, 297). She likewise says that, while her theory does take there to be a point at which sentences' value will go from 1 to less than 1 (and is therefore a 'complete categorisation' view), this aspect of the theory is not intended to get things correct; she takes drawing a line here to be a necessary consequence of capturing the features of vagueness that she's interested in (that is, vagueness as an exemplification of there being degrees of closeness to clear truth) (1997, 308-9).

The theory of vagueness defended in Smith (2008) is a little more complicated to classify, but it should be instructive to do so. Smith defends what he calls 'fuzzy plurivaluationism'. This view borrows resources from the kind of degree theory just outlined, as well as some from supervaluationism. The resource it borrows from the latter is the notion of an 'admissible interpretation' of a language. Such an interpretation is a way of assigning truth-values to the sentences of the language. Which ones count as 'admissible' is constrained by certain facts about the way language users apply the terms in that language (2008, 99); to give some easy examples, an interpretation of the language in which 'Granny Smiths are red' is given value 1 is not going to admissible, while an interpretation in which 'Dwayne "The Rock" Johnson is bald' is given value 1 is going to be admissible. Now on Smith's view each of these interpretations is an assignment of *degrees of truth* to sentences (2008, 99) (this is the 'fuzzy' aspect of it) — when we're dealing with vague terms, each such interpretation will formally look similar to an implementation of a non-plurivaluationist degree theory of vagueness such as Edgington's.[10] Plurivaluationism contrasts with supervaluationism in that, where supervaluationists 'aggregate' the values assigned on the different admissible interpretations to assess a sentence's semantic value (and so say, for example, that a sentence's *truth* amounts to it being assigned 'true' on all admissible interpretations), plurivaluationists say that these aggregated values have no semantic significance. Under a plurivaluationist theory, sentences cannot be univocally true, false (or otherwise) — rather they are true or false (or otherwise) just on some interpretations, and this is just as much the case when a sentence has some particular value under all admissible interpretations (2008, 101-2).

How should we classify this? One plausible reading is this. To the extent that Smith's plurivaluationism refuses to assign any sentences with any absolute truth-values, it can actually be said to be an extreme form of semantic quietism — on his view, no sentence can be said to be true, false, or anything else. It's also true that each admissible interpretation on Smith's view is going to represent its own implementation of a complete categorisation view, but none of these interpretations is intended to be *the* interpretation, so it would be a mistake to call plurivaluationism a complete categorisation view on this basis.

It seems that Smith's view should also be classified as an implementation of the modelling approach to logic and semantics. While his view is that there's no way to aggregate the truth-values assigned on different admissible interpretations, it's left implicit in Smith's view that, when you have a set of sentences saying of each of a set of gradually changing objects is steep, say, each of these sentences either has truth-value 1 on all admissible interpretations or does not, and as such there's a sharp boundary between the ones which do and the ones which don't. But given this aspect of Smith's theory, the theory as a whole presumably isn't aiming to get everything right about

[10]Though it won't look exactly the same — logical connectives work differently in Smith's and Edgington's theories, for example.

the semantics of vague terms — even if this kind of sharp boundary isn't supposed to represent anything semantically significant, there still doesn't seem to be any such boundary marked by any aspect of use.

To finish this section, I'd like to deliver on some promises I made early on in the chapter. I said that if we accept the claim that there is no correct semantics for vague terms, we can still endorse classical logic, without having to affirm any of the claims particular to epistemicism (under which there are complete categorisations actually drawn by meaning, which are unknowable). The above discussion should have shown how this is so — having ruled out the descriptive approach, and assuming that the two remaining approaches to logic and semantics that I've just talked about (that is, the 'prescriptive' and 'modelling' approaches) are the ones we could adopt, it's entirely possible to endorse classical logic — in the sense that's relevant to our aims in theorising about logic and semantics in the first place, whatever those aims might be — under either of these approaches. So if maintaining classical logic at all costs was a motivation for accepting epistemicism, we can see how this motivation can be undercut — we can keep classical logic without it.

I also said that, under the assumption that there is no correct semantics for vague terms, we could re-interpret Eklund's suggestion that the correct logic is the one that most closely reflects actual use, in light of the impossibility of there being a correct logic which entirely reflects actual use. Our re-interpretation should be to say: whatever the right overall approach to logic and semantics is, this approach can be best implemented by most closely reflecting actual use. This is certainly not inconsistent with either the claim that there is no correct semantics for vague terms, nor does it even force us to pick a particular approach to logic and semantics, since, as I noted, all of the approaches I've considered make room for this in being adaptable to different theories of vagueness. To clarify, though, nothing I've said previously *commits* us to

Eklund's claim in its re-interpreted form — it's just that it *can* be re-interpreted naturally to be made consistent with the different approaches I've outlined.[11]

4 Problems of Vagueness

As far as theorising about vagueness goes, if you think I'm right about what a theory of vagueness should look like and the ways in which it's going to be inherently limited, you might feel that we are pulled in two different directions. On the one hand, we want to maintain that theories of vagueness (and indeed ones that are being promoted at the time of writing) are valuable, and can represent some aspects of the phenomenon effectively, and yet it also seems that, ultimately, any semantics and/or logic of vague terms is not going to represent these terms as vague in the fullest sense of the term. An important question we face, then, is: can theories of vagueness, whether drawing complete categorisations or semantically quietist, and with whichever aim we take a semantic theory to have, actually solve the interesting problems associated with vagueness?

Let's first get an idea of the kinds of problem that theories of vagueness try to address. We can take the lead here from Keefe in identifying these problems. One task of a theory of vagueness is to give an account of the effect that vagueness has on the semantic structure of vague languages. It should also give an account of the logical relations that hold between sentences in those languages (given that they're vague), as well as explain how the truth-values of complex sentences are built up from simpler ones (again, in light of their being vague). In doing these we should specify the range of semantic statuses that sentences can have, and say how these statuses are distributed to sentences saying of different objects that they satisfy particular vague

[11] Again, the only thing that this re-interpretation misses out is the possibility that the 'pure' descriptive view of logic is correct.

terms. All of this should allow theorists to offer a general solution to sorites paradoxes, and address the phenomenon of higher-order vagueness by either denying that there is such a thing or by describing its structure (2000, 37-8).

The sorites paradox seems to rely on intuitive claims and rules of inference, and yet generates results that we want to avoid. We should therefore expect any theory which can address the problems of vagueness to involve some claims which are counter-intuitive. Keefe says that a further task for any such theory is to explain why we should accept these claims after all, and why we found the claims they contrast with intuitive to begin with (2000, 20). There are more problems that we might hope to address with a theory of vagueness, but it seems that any theory will be on good footing if it can address the ones outlined above.

So let's make a start. The literature on vagueness contains a wide variety of approaches to describing the effects of vagueness in languages on the semantic and logical structure of those languages. With appropriate modifications, it seems that any of these could be brought in line with either the requirement that it ultimately draws some complete classification, or in line with quietism — and I outlined how this could be done earlier. Any of these theories could thereby give an account of the semantic and logical structure of vague languages, or rather, the best description or representation of these languages, or a prescription for how we should reason in a vagueness-free language (depending on what view on logic and semantics we take).

We can approach the question of how complex truth-values are built up from less complex ones again by thinking about adapted versions of currently held theories. We could endorse the position of any such theory on how logical connectives should be defined. So, to give some examples, a degree theory may tell us that '$A \wedge B$' has the least truth-value of A and B, and a supervaluationist theory may say that it is not the case that if '$A \vee B$' is true, then either A is true or B is true.

We can use theories' predictions about how truth-values are built up to decide which, if any, of these provides the most accurate models of vague natural languages once sharp boundaries have been introduced (if that's what we're interested in finding out). So we might favour a (modified) degree theory less for assigning '$Fa \wedge \neg Fa$', a contradiction, value 0.5 when 'Fa' has value 0.5.[12] And we might likewise object to a (modified) supervaluationism for going against the claim that if a disjunction is true, one of its disjuncts must be true. Obviously defenders of these theories can offer replies to these objections, but I bring this up to highlight the kinds of considerations that are relevant in deciding between theories, having acknowledged that none of them will represent the correct semantics for their target language. The crucial point here is that we can theorise about complex truth-values in much the same way as we have been doing — we just have a particular set of background beliefs about the possibility of a correct semantics when we do so.

How do we solve the sorites paradox? A sorites argument for 'heap', to give an example, is going to start from two assumptions — that some particular object is a heap, and that removing just one grain of sand can't make a heap into a non-heap — and infer from them that one grain of sand must therefore be a heap. The first of these two assumptions is clearly correct, yet it's also clearly correct to say that one grain of sand is not a heap. But our claim that there is no correct semantics for 'heap' gives us a reason to say that the second assumption that the argument makes use of (the 'major' premise of the argument) cannot be correct. Recall the shape of the argument that I presented for this last claim in the previous chapter: if there were a correct semantics for vague terms, it would vindicate the assumption that there was no way to categorise all of the relevant cases into 'applies', 'doesn't apply', and 'other', but there is no semantics which vindicates this assumption which can also be said to

[12]This is a common objection to degree theories. Williamson, for example, calls this fact 'disturbing' (1994, 118).

be correct. It's just a consequence of this that the assumption that there is no way to categorise all of the relevant cases into 'applies', 'doesn't apply', and 'other' cannot be said to be correct. But if we accept that this claim can't be correct, it's difficult to see how we could also take the major premise of the sorites argument outlined above to be correct. For if it isn't correct to say that there's no way to categorise all cases into 'applies', 'doesn't apply', and 'other', then it also can't be correct to say that there's no way to categorise them all into either 'applies' or 'doesn't apply'. And it's difficult to see how we could maintain both that it's not correct to say that there's no such categorisation, *and* (our major premise) that (it's correct to say that) removing a grain of sand from a heap could not make it into a non-heap. This, in turn, serves to undermine the correctness of the conclusion of any sorites argument of this sort.

In saying that the major premise of sorites arguments of the kind described above are not correct, are we thereby committed to the correctness of its *negation*? No. The point to emphasise here is that, as I mentioned before, if there *were* a correct semantics for vague terms, it would not categorise all cases into 'applies', 'doesn't apply' or 'other', and this is not compatible with the negation of the major premise, so there's no reason to think that this negation is correct. We therefore shouldn't take either the major premise or its negation to be correct.

The kind of response to the 'standard' sorites paradox can applied just as well to any 'iteration' of it that we could come up against when thinking about higher-order vagueness. We can avoid the correctness of a sorites argument using the assumptions that there are *definite* heaps, and that removing a grain of sand couldn't make a definite heap into a non-definite heap, in much the same way, for example. The major premise of any such argument is going to rely on the assumption that (it's part of the correct semantics for 'heap' that) there's no way to categorise all of the relevant cases into those to which 'heap' applies, those to which it doesn't, and any others, since it seems difficult to reconcile the incorrectness of this assumption with the correctness of there being no point at which cases would go from ones where 'heap' applies (that is, the definite heaps) to ones where it either does not apply or which have some other status with respect to 'heap' (that is, the non-definite cases of 'heap'). And again, the assumption that the correct semantics draws no such division is an assumption that we're rejecting here.

Now what if a sorites paradox is posed to a theory of vagueness on the understanding that that theory does not intend to deliver a correct semantics for vague terms — that is, when it's the kind of theory that I've been outlining options for in this chapter? How a theory is going to respond to this challenge is going to vary between particular theories, and there is no shortage of solutions to the sorites paradox which could be adapted. Broadly speaking, though, when it comes to theories which draw complete categorisations, their solutions to the sorites paradox will be to accept that, at some level, a small change can make the difference — whether between a heap and a non-heap or an unambiguous non-heap and something else. Their apology for this need only be that their semantic theory is not intended to get everything correct (and therefore doesn't reflect the assumption that there is no way to divide up cases into 'applies', 'does not apply', and 'other'), and that having a theory of vagueness at all comes at the cost of a 'correct' resolution of the sorites paradoxes (and, again, its variants). Quietist theories will not need to commit themselves to this much — the question 'can a small change make the difference between an unambigious heap (say) and something with a different status with respect to 'heap'?' can be met by silence (or by some response that's supposed to be as good as silence); the theory may just not offer an answer because it says nothing about the status of some objects. From their perspective, though, this isn't avoiding any serious question, or failing to acknowledge something correct in sorites arguments, since they take there to be nothing that they can say about the cases they fall silent about altogether.

Is there a 'use' version of the sorites paradox left over? It seems not, since it seems that competent speakers neither should, nor in fact do, make sets of judgements that reflect the sorites paradox, even when prompted. The pattern of reasoning involved in the sorites paradox is difficult to find clear flaws in (even after giving it a decent amount of thought), yet it doesn't seem to be a feature of the use of vague terms that competent users (other than a minority of theorists of vagueness) will actually accept the contradictory conclusion of sorites arguments. When pressed on particular cases, it's more reasonable to expect a diversity of responses, but ultimately for speakers to 'draw lines' somewhere, even if this just means refusing to say whether some shades are green, or expressing doubt, or indeed to claim to have identified the last of the relevant kind. The question of how to make sense of this diverse set of responses is interesting, but it's not the sorites paradox.

Why did we find the sorites paradox and its variants so enticing to begin with? We can begin to offer an explanation on the basis of what I've said so far. For one, it is true *in a sense* that there's no grain of sand that makes the difference to whether something is a heap or not (or to whether something is a heap or has some other status with respect to 'heap') — the sense in which competent speakers' use doesn't mark any such difference (I argued for this in chapter 4). It's also true that if there were a correct semantics for vague terms, it would reflect this feature of use: the correct semantics of vague terms, if there were one, would make it true that no grain of sand makes the difference to whether something was a heap or not (or, again, to whether something was a heap or had some other status). It's entirely natural, then, to feel from our linguistic competence that we can't definitively make assertions which identify one grain of sand that makes the difference, and to read from this feeling that there is no such grain. If I'm correct in my conclusions this far, the mistake in making this last inference is quite subtle, in that it relies on the obvious-sounding assumption that there is a correct semantics for vague terms, which accurately reflects the way they're used.

To draw things to a close, let's come back to the topic of higher-order vagueness. How do the theories of vagueness that I'm envisioning deal with this phenomenon? For one, as I pointed out in the previous chapter, it seems that they should say that if a term, say 'red', is vague, then '(epistemic) borderline case of 'red'' satisfies a necessary condition for vagueness, and so does '(epistemic) borderline case of 'borderline red'', and so on. That is, the competent use of 'red' doesn't draw any consistent categorisations of objects which are red, those which aren't, and any others; and that competent use of 'borderline red' doesn't draw any such categories, and competent use of 'borderline borderline red' doesn't draw any, and so on. I also pointed out that the same goes for semantic interpretations of 'borderline red', 'borderline borderline red', and so on.

It should be no surprise that a consequence of what I've suggested so far is that, just as there'll be no correct semantics for 'red' (assuming that 'red' is vague), there'll likewise be no correct semantics for 'borderline red', 'borderline borderline red', and so on. Our treatment of these terms should therefore follow the same pattern of treatment that terms such as 'red' have received so far. We could treat any such term using either a 'complete categorisation' approach or a quietist approach, the details of which could be spelled out in any number of different ways. A theory on which complete categorisations are given could take the form of either of the two kinds of theory of higher-order vagueness that I outlined earlier — columnar or non-columnar. In the latter case, this would require some adaptation: as I argued in chapter 2, non-columnar views can't be made consistent if they take there to be at least one of each of an infinite number of borderline categories, and so such a theory would need to artificially 'cut off' the higher-order vagueness at some point when we're looking at finite sets of objects. Quietist approaches can't be columnar in Bobzien's sense of dividing up all of the cases into three positive kinds — that is, the clear positive cases, the borderline cases (at all orders) and the clear negative cases, since this would just be to assign a significant status to all cases, but they could adopt columnar views in the sense of sharply categorising cases into clear positive cases, borderline cases (at all orders), clear negative cases, and cases about which it goes silent, as well as non-columnar views of higher-order vagueness in a variety of ways. Again, this can't involve all borderline categories of an infinite number being represented over a finite set of objects, though rather than drawing an explicit cut-off somewhere, a quietist theory could simply be silent about the existence of borderline cases over a certain order.

There still seem to be some big questions left over, that we might want even a theory that's not trying to get things correct to answer. For one, we might think that a theory of vagueness should say something about some of the characteristic features of vagueness that we put to one side in chapter 4. In particular, we might ask whether vague terms are epistemically tolerant (that is, whether it's possible to know that they apply in some cases and know that they don't apply in marginally different cases), and whether all vague terms have borderline cases (and indeed whether the epistemic tolerance of vague terms entails that they have borderline cases, or vice versa). In turn, these questions might lead to the following questions: Do 'red', 'borderline red' and 'borderline borderline red', and so on, all have borderline cases? If so, how are these different categories of borderline case related to one another? How are they structured? Doesn't this lead back to higher-order vagueness paradoxes, whatever our answer is?

The position we should take on this — which is admittedly slightly unsatisfying — is that there are not going to be full, correct answers to all of these questions. If we could say, once and for all, which kinds of borderline cases existed, and exactly what structure they had, for one, we would in effect be putting forward a complete categorisation of cases as *correct*, which is something we're trying to avoid. This isn't to say that there aren't correct things that we can say which would in part answer some of these questions — for one, it seems likely to me that 'red' had borderline cases, and that borderline cases of 'red' aren't definite non-borderline cases of 'red'; that seems to be an obvious structural fact about borderline cases at different orders — it's just that a complete theory is not on the table, and it seems that this is what these questions are really asking for. I'd also like to stress here that a theory which wasn't aiming to give the correct semantics *could* give full answers to these questions — they just won't be correct.

The question about epistemic tolerance complicates this further, as the argument presented in chapter 1 from Greenough suggests that if all vague terms are epistemically tolerant, then they all have borderline cases. And if we admit, for example, that when a term like 'red' is vague, 'borderline red' and 'borderline borderline red' and so on are also vague, then there's an argument looming which is going to tell us that (it's *correct* that) there are some objects in each borderline category. There are a few things to say about this. First, it's worth noting that the argument that Greenough offers (which I outlined in more detail in chapter 1) relies on classical logic — in particular it works by offering an argument similar in structure to the sorites paradox itself (relying on looking at what's knowable about successive objects), and making essential use of the reductio and double negation elimination rules. This is reason enough to be sceptical of Greenough's argument in itself, since a correct logic that assigns a semantic status (that is, 'applies', 'doesn't apply', or some 'other' status) to each of a succession of cases involved in this argument (and what can be known about those cases) is one thing I've been arguing against, and which is exactly a consequence of classical logic.

Second, we *could* take all vague terms to be (correctly described as) epistemically tolerant, and likewise to all have borderline cases, and we could in turn take (for example) 'borderline red', 'borderline borderline red', and so on, to all have borderline cases. One way of doing this consistently would be to borrow partially from the columnar views we saw in chapter 3 and say that there's some object which was a borderline case of 'red', a borderline borderline case, a borderline borderline borderline case, and so on. It's consistent, at least, to do this because it would not amount to drawing complete categorisations of cases — it amounts to saying a lot about one case in particular. But committing to epistemic tolerance and borderline cases as necessary features of vague terms (and in turn higher-order borderline vagueness) leaves us very constrained in what we can say about how these features are exhibited. We couldn't take a columnar view along the lines of Bobzien's view from chapter 3 to be correct, since this view effectively divides up all cases into three categories (borderline cases at all orders, definite positives at all orders, and definite negatives at all orders). We also couldn't endorse a non-columnar view as correct (or, at least, one that doesn't 'lapse' into columns eventually) — as the arguments I outlined in chapter 2 showed, non-columnar views of higher-order vagueness end up making the borderline cases at successive orders take up more and more 'space' within the finite orderings of objects, ultimately leaving insufficient room for borderline cases to actually be structured in a non-columnar way. The alternative to all of this is to say that there's no correct way of saying whether all vague terms are epistemically tolerant, or admit of borderline cases (or exhibit higher-order *borderline* vagueness specifically). Whichever way we go here, my point is that there are options open which allow us to accommodate a number of different claims we might want to make about vagueness; we just need to make choices about which ones we think are most important to include.

5 Conclusion

This chapter has provided an outline of some of the consequences of the claim I argued for in chapter 5, that there is no correct semantics for vagueness. To sum up, this claim does not require us to make huge revisions to the currently popular theories of vagueness (or indeed to go about theorising about vagueness in radically different ways), but it *does* point to some limitations on what these theories can do. In particular it shows that they cannot give a completely correct semantics for vague terms (though this can be mitigated in different ways, depending on our aims), and that they cannot give a correct treatment of the phenomenon of higher-order vagueness.

To bring this thesis to a close, let's reflect on some of the key results that we've established. In chapters 2 and 3 we saw that there are significant reasons to reject the orthodox approach to characterising vagueness — in terms of 'no sharp boundaries' (understood epistemically or otherwise), borderline cases, and higher-order borderline vagueness. Chapter 2 specifically focused on the threat to this characterisation from paradoxes of higher-order vagueness, showing that we can we can generate these paradoxes using progressively fewer resources, and indeed that we can generate one such paradox using just the assumption that there is higher-order vagueness which is *non-columnar* in structure. Chapter 3 showed that retreating from there to a columnar view of higher-order vagueness is not a promising position by showing that Bobzien's columnar view, the only attempt in the literature to spell out such a view, could only be made to work on an implausible picture of ideal language use, and even then contains a serious inconsistency.

Having rejected the orthodox characterisation of vagueness, we then settled on a new one in chapter 4, drawing from related suggestions from Eklund and Horgan, giving the following necessary condition on a term's being vague: if a term is vague, there could be a set of objects gradually changing with respect to that term such that competent users of that term wouldn't make judgements which stably categorised those objects into those to which the term applies, those to which it does not, and any others, even after exhausting any investigation that they could carry out. In chapter 5 I presented an argument showing that there is no correct semantics for vague terms.

This proceeded by showing what such a correct semantics would need to look like, given the necessary condition established in the previous chapter, but then showed how no semantics could reflect this requirement, going through arguments from Fine, Horgan and Sainsbury before settling on my own as an argument form which could be applied to any theory claiming to provide a correct semantics, showing that any such theory either lapses into quietism or a complete categorisation theory. This final chapter has dealt with the fallout of this result. Here we've sketched out four kinds of theory of vagueness which could be adopted in spite of that result — combining 'complete categorisation' and 'quietist' approaches to theorising about vagueness specifically, on the one hand, with either a 'prescriptive' or 'modelling' approach to logic and semantics more generally. On the way, we also saw that a third approach to logic and semantics — the 'descriptive' approach — is actually incompatible with my claim that there is no correct semantics for vague terms, which I suggest is a reason to reject this approach. Having shown that, I then pointed the way forwards on how we might go about solving a variety of the traditional problems of vagueness, showing how the theoretical revisions I've been proposing can actually help to do so.

Bibliography

S. Alxatib and F. J. Pelletier. The Psychology of Vagueness: Borderline Cases and Contradictions. *Mind and Language*, 26(3):287–326, 2011.

S. Bobzien. Higher-order Vagueness, Radical Unclarity, and Absolute Agnosticism. *Philosophers' Imprint*, 10(10):1–30, 2010.

S. Bobzien. In Defense of True Higher-order Vagueness. *Synthese*, 180(3):317–335, 2011a.

S. Bobzien. If It's Clear, Then It's Clear That It's Clear, or is It? Higher-order Vagueness and the S4 Axiom. In B. Morison and K. Ierodiakonou, editors, *Episteme, etc.* Oxford University Press, 2011b.

S. Bobzien. Higher-order Vagueness and Borderline Nestings: A Persistent Confusion. *Analytic Philosophy*, 54(1):1–43, 2013.

S. Bobzien. Higher-order Vagueness and Numbers of Distinct Modalities. *Disputatio*, VI(39):131–137, 2014.

S. Bobzien. I—Columnar Higher-Order Vagueness, or Vagueness is Higher-Order Vagueness. *Aristotelian Society Supplementary Volume*, 89(1):61–87, 2015.

S. Bobzien. Refutation of the Objection that Columnar Higher-order Vagueness is Incoherent, Draft. URL http://www.academia.edu/8093703/Refutation_of_the_objection_that_columnar_higher-order_vagueness_is_incoherent. Accessed 15th September 2018.

J. Cargile. The Sorites Paradox. *British Journal for the Philosophy of Science*, 20(3):193–202, 1969.

B. F. Chellas. *Modal Logic: An Introduction*. Cambridge University Press, 1980.

D. Davidson. *Inquiries Into Truth And Interpretation*. Oxford University Press, 1984.

D. Edgington. Wright and Sainsbury on Higher-order Vagueness. *Analysis*, 53(4): 193–200, 1993.

D. Edgington. Vagueness by Degrees. In R. Keefe and P. Smith, editors, *Vagueness: A Reader*. MIT Press, 1997.

M. Eklund. Inconsistent Languages. *Philosophy and Phenomenological Research*, 64(2): 251–75, 2002.

M. Eklund. What Vagueness Consists In. *Philosophical Studies*, 125(1):27–60, 2005.

D. G. Fara. Shifting Sands: An Interest Relative Theory of Vagueness. *Philosophical Topics*, 28(1):45–81, 2000. Originally published under the name "Delia Graff".

D. G. Fara. Gap Principles, Penumbral Consequence, and Infinitely Higher-order Vagueness. In J. C. Beall, editor, *Liars and Heaps: New Essays on Paradox*. Clarendon Press, 2004. Originally published under the name "Delia Graff".

K. Fine. Vagueness, Truth and Logic. *Synthese*, 30(3-4):265–300, 1975.

K. Fine. The Impossibility of Vagueness. *Philosophical Perspectives*, 22(1):111–136, 2008.

G. Frege. *Begriffsschrift: Eine Der Arithmetische Nachgebildete Formelsprache des Reinen Denkens*. L. Nebert, 1879.

M. Glanzberg. Against Truth-Value Gaps. In J. C. Beall, editor, *Liars and Heaps: New Essays on Paradox*. Clarendon Press, 2004.

P. Greenough. Vagueness: A Minimal Theory. *Mind*, 112(446):235–281, 2003.

P. Greenough. On What It is to Be in a Quandary. *Synthese*, 171(3):399–408, 2009.

R. Heck. A Note on the Logic of (Higher-order) Vagueness. *Analysis*, 53(4):201–208, 1993.

T. Horgan. Robust Vagueness and the Forced-March Sorites Paradox. *Philosophical Perspectives*, 8(Logic and Language):159–188, 1994.

T. Horgan. Transvaluationism: A Dionysian Approach to Vagueness. *Southern Journal of Philosophy*, 33(S1):97–126, 1995.

T. Horgan. The Transvaluationist Conception of Vagueness. *The Monist*, 81(2):313–330, 1998.

T. Horgan. Facing Up to the Sorites Paradox. *The Proceedings of the Twentieth World Congress of Philosophy*, 6:99–111, 2000.

T. Horgan. Transvaluationism. *The Harvard Review of Philosophy*, 14(1):97–125, 2006.

T. Horgan. Transvaluationism About Vagueness: A Progress Report. *Southern Journal of Philosophy*, 48(1):67–94, 2010.

G. E. Hughes and M. Cresswell. *A New Introduction to Modal Logic*. Routledge, 1996.

D. Hyde. Why Higher-order Vagueness is a Pseudo-Problem. *Mind*, 103(409):35–41, 1994.

D. Hyde. From Heaps and Gaps to Heaps of Gluts. *Mind*, 106(424):641–660, 1997.

D. Hyde. The Sorites Paradox. In G. Ronzitti, editor, *Vagueness: A Guide*, pages 1–17. Springer Verlag, 2011.

D. Hyde and M. Colyvan. Paraconsistent Vagueness: Why Not? *Australasian Journal of Logic*, 6:107–121, 2008.

R. Keefe. *Theories of Vagueness*. Cambridge University Press, 2000.

R. Keefe. II—Modelling Higher-Order Vagueness: Columns, Borderlines and Boundaries. *Aristotelian Society Supplementary Volume*, 89(1):89–108, 2015.

W. G. Lycan. *Logical Form in Natural Language*. MIT Press, 1984.

A. Mahtani. Thinking Precisely about Vagueness: an Interview with Anna Mahtani, 2014. URL `http://eprints.lse.ac.uk/80857/`. Accessed 15th September 2018.

R. Mena. The Quietist's Gambit. *Crítica: Revista Hispanoamericana de Filosofía*, Forthcoming.

R. Montague. *Formal Philosophy; Selected Papers of Richard Montague*. New Haven: Yale University Press, 1974.

H. Montgomery and R. Routley. Contingency and Non-Contingency Bases for Normal Modal Logics. *Logique Et Analyse*, 9(35):318, 1966.

G. Priest. Inclosures, Vagueness, and Self-Reference. *Notre Dame Journal of Formal Logic*, 51(1):69–84, 2010.

H. Putnam. Vagueness and Alternative Logic. *Erkenntnis*, 19(1-3):297–314, 1983.

W. V. Quine. The Scope and Language of Science. *British Journal for the Philosophy of Science*, 8(29):1–17, 1957.

D. Raffman. Borderline Cases and Bivalence. *Philosophical Review*, 114(1):1–31, 2005.

D. Raffman. *Unruly Words: A Study of Vague Language*. Oxford University Press, 2014.

M. Sainsbury. Is There Higher-order Vagueness? *Philosophical Quarterly*, 41(163):167–182, 1991.

M. Sainsbury. Sorites Paradoxes and the Transition Question. *Philosophical Papers*, 21 (3):177–190, 1992.

R. M. Sainsbury. Concepts Without Boundaries, 1990. King's College London Department of Philosophy. Inaugural lecture, delivered 6th November 1990.

S. Shapiro. I—Context, Conversation, and So-Called 'Higher-order Vagueness'. *Aristotelian Society Supplementary Volume*, 79(1):147–165, 2005.

S. Shapiro. *Varieties of Logic*. OUP Oxford, 2014.

N. J. J. Smith. *Vagueness and Degrees of Truth*. Oxford University Press, 2008.

R. A. Sorensen. An Argument for the Vagueness of 'Vague'. *Analysis*, 45(3):134–137, 1985.

R. A. Sorensen. *Vagueness and Contradiction*. Oxford University Press, 2001.

A. Tarski. The Concept of Truth in Formalized Languages. *Logic, Semantics, Metamathematics*, 2:152–278, 1956.

M. Tye. Vague Objects. *Mind*, 99(396):535, 1990.

M. Tye. Sorites Paradoxes and the Semantics of Vagueness. *Philosophical Perspectives*, 8:189–206, 1994.

A. C. Varzi. Super-duper Supervaluationism. In T. Childers and O. Majer, editors, *Logica '94. Proceedings of the 8th International Symposium*. Filosofia. Prague, 1995.

A. C. Varzi. Inconsistency Without Contradiction. *Notre Dame Journal of Formal Logic*, 38(4):621–639, 1997.

A. C. Varzi. Higher-order Vagueness and the Vagueness of 'Vague'. *Mind*, 112(446): 295–298, 2003.

B. Weatherson. Vagueness as Indeterminacy. In R. Dietz and S. Moruzzi, editors, *Cuts and Clouds: Vaguenesss, its Nature and its Logic*. Oxford University Press, 2010.

T. Williamson. *Vagueness*. Routledge, 1994.

T. Williamson. Imagination, Stipulation and Vagueness. *Philosophical Issues*, 8:215–228, 1997.

T. Williamson. On the Structure of Higher-order Vagueness. *Mind*, 108(429):127–143, 1999.

T. Williamson. Horgan on Vagueness. *Grazer Philosophische Studien*, 63(1):273–285, 2002.

C. Wright. On the Coherence of Vague Predicates. *Synthese*, 30(3-4):325–65, 1975.

C. Wright. Is Higher-order Vagueness Coherent? *Analysis*, 52(3):129–139, 1992.

C. Wright. On Being in a Quandary. *Mind*, 110(1):45–98, 2001.

C. Wright. Vagueness: A Fifth Column Approach. In J. C. Beall, editor, *Liars and Heaps: New Essays on Paradox*. Oxford University Press, 2003.

C. Wright. The Illusion of Higher-order Vagueness. In R. Dietz and S. Moruzzi, editors, *Cuts and Clouds. Vaguenesss, its Nature and its Logic*. Oxford University Press, 2009.

E. Zardini. A Model of Tolerance. *Studia Logica*, 90(3):337–368, 2008.

E. Zardini. Higher-order Sorites Paradox. *Journal of Philosophical Logic*, 42(1):25–48, 2013.

E. Zardini. First-Order Tolerant Logics. *Review of Symbolic Logic*, Forthcoming.