Too Many Cats:

The Problem of the Many and the Metaphysics of Vagueness

Nicholas K. Jones

Ph.D. Thesis

Birkbeck, University of London (School of Philosophy)

Submitted September 2010

Abstract

Unger's Problem of the Many seems to show that the familiar macroscopic world is much stranger than it appears. From plausible theses about the boundaries of ordinary objects, Unger drew the conclusion that wherever there seems to be just one cat, cloud, table, human, or thinker, really there are many millions; and likewise for any other familiar kind of individual. In Lewis's hands, this puzzle was subtly altered by an appeal to vagueness or indeterminacy about the the boundaries of ordinary objects. This thesis examines the relation between these puzzles, and also to the phenomenon of vagueness.

Chapter 1 begins by distinguishing Unger's puzzle of too many candidates from Lewis's puzzle of borderline, or vague, candidates. We show that, contra Unger, the question of whether this is a genuine, as opposed to merely apparent, distinction cannot be settled without investigation into the nature of vagueness. Chapter 2 begins this investigation by developing a broadly supervaluationist account of vagueness that is immune to the standard objections. This account is applied to Unger's and Lewis's puzzles in chapters 3 and 4. Chapter 3 shows that, despite its popularity, Lewis's own approach to the puzzles is unsatisfactory: it does not so much solve the puzzle, as prevent us from expressing them; it cannot be extended to objects that self-refer; it is committed to objectionable theses about temporal and modal metaphysics and semantics. Chapter 4 develops a conception of ordinary objects that emphasises the role of identity conditions and change, and uses it to resolve both Problems of the Many. This allows us to diagnose the source of the puzzles: an overemphasis on mereology in contemporary material ontology.

I hereby declare that all the work presented in this thesis is my own:

Nicholas K. Jones

In loving memory of Richard Blundell

Acknowledgements

I have benefited greatly from the wisdom of my supervisors at Birkbeck: Dorothy Edgington, Keith Hossack, Fraser MacBride and Ian Rumfitt all discussed drafts (or ancestors of drafts) of various chapters. Without their feedback and encouragement this thesis would certainly have been much worse than it actually is. Of everyone I've met since coming to Birkbeck, Dorothy and Fraser deserve special thanks; their assistance at every stage of this project went far beyond that which I had any right to expect.

Innumerable conversations with Will Bynoe have shaped many of the views here; I'm grateful for his good sense and constant willingness to talk philosophy (as well as film, science-fiction, boardgames and comics). The Birkbeck graduate community provided a friendly environment in which to learn from my mistakes; conversations with Alex Douglas, Simon Hewitt, Sam Lebens, Steven Methven, Gina Tsang and Dan Turnbull stand out as particularly enjoyable and constructive. Thanks to Mark Textor for feedback on part of chapter 2, and to the King's Metaphysics Society for providing an audience for early versions of many of the arguments here. Thanks also to Rory Madden, Jonny McIntosh and Lee Walters.

I received funding from the Arts and Humanities Research Council during my first three years at Birkbeck, and a Royal Institute of Philosophy Jacobsen Fellowship whilst I was writing-up; thanks to both organisations.

I count myself lucky for the unceasing support of all my friends and family. I'm especially grateful to Lara Freeman for competing with philosophy and always having faith in me, as well as for putting up with a sometimes thoughtless and distracted boyfriend. My greatest debt is to my parents, for everything they've ever done for me.

Contents

Analytical Contents

Introduction

Unger's Problem of the Many seems to show that almost all of our ordinary numerical judgements are radically mistaken: for any kind K of ordinary macroscopic object, wherever there appears to be just one K, really there are many millions. Lewis presented a similar argument, though he appealed to vagueness or indeterminacy, whilst Unger did not. This thesis investigates the relation between these puzzles, and their connection to the phenomenon of vagueness.

Chapter 1 develops versions of the puzzles that require only very minimal assumptions. We show the puzzles do not primarily concern the existence of individuals, but the instantiation of ordinary sortal properties. Puzzles about the number of objects arise only via the commonplace assumption that each instantiation of an ordinary sortal is by a single individual; even those who deny that ordinary composite individuals exist must address the Problem of the Many. We will also see that, contra Unger, the question of whether he and Lewis were addressing the same puzzle cannot be settled prior to an investigation into the nature of vagueness. The chapter closes by arguing that there are good reasons to reject Unger's conclusion: our ordinary numerical judgements are not typically radically mistaken. The conclusion of Unger's puzzle is not merely implausible—surely there are not many humans seated in my chair, writing this thesis—but creates significant problems in the metaphysics of time, modality, free will, choice, moral responsibility and singular thought.

Chapter 2 turns to vagueness, and supervaluationist theories of vagueness in particular. We develop two broadly supervaluationist accounts of vagueness. One treats vagueness as a semantic phenomenon, and uses classes of sharpenings to represent vague semantic structures. The other treats vagueness as a partly semantic and partly metasemantic phenomenon; classes of sharpenings are used to represent classes of semantic structures that fit the meaning-determining facts well enough to count as the actual, or intended, semantic structures of a vague language. On this second view, vague languages express many precise contents. We argue that the standard objections to supervaluationism do not touch the second view, and hence that it is preferable to the first. That view is applied to the Problem of the Many in chapters 3 and 4. Due to restrictions on space, a study of a recent argument for the classicality of supervaluationist logic, due to J.R.G. Williams, has been removed, and is now forthcoming in the *Journal of Philosophy*.

Chapter 3 serves two purposes. The first is an examination of Lewis's solution to his puzzle. Lewis claimed that his and Unger's puzzles are sources of referential vagueness in our names for ordinary objects, and used the supervaluationist technique to ensure that sentences that express our ordinary numerical judgements are true. We show that this approach: does not solve the problems, but merely prevents us from expressing them; cannot be extended to self-referrers; is committed to objectionable theses about the metaphysics and semantics of temporal and modal discourse. The second goal of the chapter is to defend the supervaluationist view developed in chapter 2 against objections due to Schiffer, Barnett, McGee and McLaughlin, and Sorensen. These objections all concern supervaluationist accounts of vague reference.

Chapter 4 closes by developing a response to Unger's puzzle and then extending it to Lewis's. We begin with the following thesis about ordinary objects: change and persistence are explanatorily prior to mereology, constitution and location. We use this thesis to argue that Unger's puzzle shows only that ordinary macroscopic objects may not have a unique collection of microscopic constituents, not that there are many such objects where there appears to be only one: a single ordinary object may be simultaneously constituted by several (partially disjoint) portions of matter. Having developed two versions of this view, we close with three ways of extending it to Lewis's puzzle of constitutional vagueness.

Chapter 1

Two Problems of the Many

In "The problem of the many", Peter Unger presented a novel and intriguing puzzle about ordinary macroscopic material objects (Unger, 1980). From plausible theses about the boundaries of these objects, Unger drew the conclusion that our ordinary numerical judgements about them are radically mistaken. David Lewis addressed this in his "Many, but almost one" (1993a). But in Lewis's hands the problem was subtly altered by an appeal to vagueness or indeterminacy in the boundaries of ordinary objects. It is therefore not clear whether Unger and Lewis were addressing the same problem. The issue turns in part upon the nature of vagueness. One goal of this thesis is clarity about the relation between these puzzles, and also the phenomenon of vagueness. Another is greater clarity about the metaphysical and semantic commitments of potential solutions to the puzzle. Yet another is to try and solve the puzzles. The first step towards achieving these goals is to find out what the puzzles are. That is the purpose of this chapter.

Unger's puzzle is presented in §1.1, and Lewis's in §1.2. We turn to their relationship in §1.3. §1.4 closes by asking whether these are genuine problems or mere puzzles. The next chapter develops a supervaluationist approach to vagueness. Different applications of this approach to the Problem of the Many are examined in chapters 3 and 4.[1]

[1] Peter Geach presented a similar puzzle in §110 of the third edition of *Reference and Generality*, which he attributed to William of Sherwood (Geach, 1980). A recent ancestor of this thesis contained an examination of the relation between Geach's puzzle and those of Unger and Lewis, as well as Geach's doctrine of Relative Identity. The result of including that discussion was however, far too

1.1 Unger's puzzle

This section presents Unger's puzzle. Although versions arise for all ordinary kinds of macroscopic object, our initial presentation and discussion follows Unger in focusing upon clouds. It is comparatively easy to see how Unger's puzzle arises for clouds, despite it being perhaps somewhat doubtful whether they are a genuine kind of individual.

The initial presentation of the puzzle is in §1.1.1. §1.1.2 clarifies the puzzle and weakens its ontological assumptions. A key premiss is examined in §1.1.3, and an alternative provided in §1.1.4. §1.1.5 closes by extending the puzzle from clouds to other ordinary sorts of object.

1.1.1 How many clouds?

Unger begins by asking us to consider a typical cloud, C. Let a be a water molecule on C's left-hand boundary;[2] let b be a water molecule in C's exterior but extremely close to its right-hand boundary. Consider the object D that differs from C only by excluding a and including b:

Unger asks: is D a cloud? He answers: yes.

Not just anything is a cloud; an appropriate internal structure and constitution are required. Let resemblance in cloud-respects be resemblance w.r.t. this structure and constitution. Unger (1980, p.122) endorses the following *principle of minute differences*:

PMD1 If x is a typical cloud and y differs only minutely in cloud-respects from x, then y is a cloud.[3]

long a thesis; so it was removed shortly prior to submission.

[2] We assume that cloud interiors are closed: droplets on a cloud's boundary are in its interior. Nothing turns on this.

[3] Initial universal quantifiers will often be omitted in the interests of readability.

Since C is a typical cloud and D differs only minutely in cloud-respects from C, it follows that D is a cloud. Since C and D each has a part, the droplets a and b respectively, that the other does not, they are distinct clouds.

Why grant that D differs only minutely in cloud-respects from C? Cloud-interiors are characterised by a high density of suspended water droplets, and cloud-exteriors by a low density thereof.[4] The transition from high density interior to low density exterior is gradual, not marked by any sharp fall in droplet-density. This ensures that we can select droplet b from sufficiently close to C's boundary to guarantee that D differs only minutely in cloud-respects from C. Since the boundaries of all typical clouds are like this, we can generalise: for any typical cloud, another cloud is almost co-located with it.

We need not stop at two clouds. Millions of droplets lie on the boundary of each typical cloud. And millions of droplets in each typical cloud's exterior are extremely close to its boundary. Any two such would suffice in place of a and b. There is one cloud for each of these pairs. So at least 10^{12} clouds are almost co-located with each typical cloud. We could even have included or excluded two (or three, or...) droplets and still obtained objects that differ minutely in cloud-respects from C. So really there will be even more clouds than this.[5]

How can this be? The intuitive view is that, sometimes, a typical cloud is the only cloud in the sky. But from this premiss, Unger's argument leads to the contradictory conclusion that, in those situations, there are billions of clouds in the sky. At best, our numerical judgements about clouds are in radical error: there are vastly many more than we thought. Unger thinks that if clouds exist, then our ordinary numerical judgements about them are typically correct. He sees this as a non-negotiable component of our ordinary world-view.[6] So, he claims, clouds

[4] We assume that only water droplets are parts of clouds. Nothing turns on this simplification.

[5] Droplets slightly further into C's interior and exterior would presumably have been acceptable too.

[6] Unger's most recent work on the puzzle rejects this (Unger, 2006a, ch.7). It is unproblematic, he thinks, for millions of clouds, tables, plants, and maybe even cats, to almost coincide where there seems to be only one. But he does regard it as non-negotiable that he is the only conscious being in his immediate vicinity, and he thinks that each of us will think the same about ourselves. He concludes that only a form of mind-body substance dualism can respect this.

cannot exist: nothing can satisfy our concept *cloud*.

1.1.2 Dispensing with fusions

The section examines and weakens the ontological assumptions of Unger's argument. Our original presentation assumed that some object D differs from C only by (i) not including one of C's boundary-droplets, and (ii) including one extremely close droplet in C's exterior. Why think that there is such an object? Does Unger's argument require it?

D's existence is plausible, if D is conceived as a lump or portion of matter. §1.1.2.1 examines this suggestion. We'll see that this creates trouble for Unger's use of PMD1: it's doubtful whether any mere portion of matter closely resembles any cat in cat-respects. Three responses to this objection will then be considered: a deviant temporal and modal semantics; a very liberal theory of matter; a modified principle of minute differences. The third is most satisfactory. But even this is committed to the existence of arbitrary lumps of matter, and that assumption is not beyond reproach. §1.1.2.2 invokes the apparatus of plural logic to show that D's existence is an inessential assumption. We thereby strengthen Unger's argument by maximally weakening its ontological assumptions.

Some terminology will aid our discussion. Using the notion of improper parthood—the sense of parthood in which everything is a part of itself—we define:

> x *overlaps* y iff something is part of both x and y.
>
> x is *disjoint* from y (also: x *excludes* y) iff x does not overlap y.
>
> x *includes* y iff everything that overlaps y also overlaps x.
>
> x is a *fusion* of set s (also: x *fuses* s; s *composes* x) iff (i) everything that overlaps x overlaps some member of s, and (ii) everything that overlaps some member of s overlaps x.

Let s_0 be the set of water droplets that composes our typical cloud C; let s_B be the set of droplets on C's boundary; let s_E be the set of droplets only just in C's exterior; let $s_1, \ldots, s_n$ be all the sets whose members are (i) every member of s_0 except some one member of s_B, and (ii) some member of s_E; let each D_i amongst $D_0, \ldots, D_n$ be a fusion of s_i. Since there are millions of droplets in s_B and s_E, there are millions of cloud-candidates D_i.

The argument for millions of clouds assumes that each s_i has a fusion D_i. The principle of minute differences PMD1 is then invoked to conclude that each D_i is a cloud. Why grant that such fusions exist? Does Unger require them?

1.1.2.1 Fusion and lumps of matter

Does every candidate D_i exist? The following entails that they do:

Unrestricted Fusion Every set has a fusion.

Although Unrestricted Fusion has its defenders, notably Lewis (1991) and Theodore Sider (2001a), it is highly controversial. Despite this controversy, there are entities for which it is plausible: lumps (or portions) of matter, masses and space-time points.[7] This section considers the following question: can we take the D_i's in Unger's argument as lumps of matter? If so, then the result is a version of the argument in which the existence of the candidates isn't overly controversial.

We begin with a problem for this account of the candidates: since no lump and cloud resemble one another closely in cloud-respects, PMD1 doesn't imply that any of the candidates is a cloud. Three kinds of response will be examined. The first invokes a non-standard temporal and modal semantics in order to block the argument against clouds and lumps resembling one another in cloud-resects. The second response invokes a more liberal theory of matter. Both these responses will be rejected. The section closes with a more satisfying response that invokes a different principle of minute differences. The next section shows how to do without the existence of individual candidates entirely.

Matter and minute differences Material objects are, in some sense, made out of matter. Unrestricted Fusion is an attractive thesis about matter. To see this, consider any collection of material objects. Some (lump of) matter makes them

[7] Throughout, we use 'lump' in a semi-technical sense for a porion of the material "stuff" from which ordinary macroscopic objects are made, whatever that stuff might ultimately turn out to be.

up.[8] That lump is a fusion of the set of those objects. So Unrestricted Fusion holds for lumps.[9] Taking each candidate D_i as the lump from which the members of the set s_i are made (and hence as a fusion of that set), allows us to appeal to this intuitively attractive argument in support of their existence.

This approach is problematic. Following are two reasons to doubt that any cloud and lump of matter resemble one another at all closely in cloud-respects. Each is a reason to doubt therefore, that the principle of minute differences PMD1 implies that any lump D_i is a cloud.

First reason: lumps have only permanent and necessary parts, while clouds do not.

> x is a *permanent part* of y iff x is part of y at every time at which y exists.
>
> x is a *necessary part* of y iff, necessarily, if y exists, then x is part y.[10]

Lump-mereology is modally and temporally invariant. Since clouds can have different droplet-parts at different times and could have had different droplet-parts than they actually do, clouds differ significantly from lumps in cloud-respects.

Second reason: clouds and lumps have different existence and identity conditions. For example, a lump exists iff its constituent sub-portions of matter exist, but clouds exist only when their droplets are sufficiently densely arranged. Since no lump and cloud have the same existence conditions, no lump closely resembles a cloud in cloud-respects.

A successful reduction of ordinary objects to lumps of matter would avoid both these problems. There are two strategies this reduction might take. The first in-

[8] Note the collective reading here: for any objects, some matter makes them up without making any one of them up (unless "they" are one).

[9] The assumption that, for any objects, some lump makes them up is essential here, and tantamount to our conclusion. The point is not to provide independent argument for Unrestricted Fusion, but merely to illustrate how natural it is for lumps.

[10] Interestingly, masses seem to be unlike lumps of matter in this respect: only some parts of masses need be permanent or necessary, specifically, those of the kind of which it is a mass. If x is a mass of water, then x's water molecule parts are permanent and necessary parts of x. But the parts of those water molecules may be neither permanent nor necessary parts of x, despite being parts of x. For water molecules can survive changes in their constituent electrons without endangering the existence of any water-mass of which they are parts. See Zimmerman (1995) for more.

vokes a non-standard temporal and modal semantics. The second invokes a more flexible theory of matter. We consider and reject these strategies in turn. We then present an alternative non-reductive strategy involving a different principle of minute differences.

First strategy: counterpart-theory The problem with taking the candidates D_i as lumps of mater was that clouds and lumps have different modal and temporal profiles. We now examine the use of counterpart-theory to block the argument from modal and temporal differences to non-identity. If successful, this will allow us to maintain that clouds and lumps satisfy exactly the same modal and temporal formulae, and hence undermine the argument for significant differences in cloud-respects between clouds and lumps.

David Lewis introduced counterpart-theory as an account of *de re* modal predication (Lewis, 1968, 1971, 1986b). Although Lewis formulated it as an extensional translation of *de re* modal discourse, not as a semantic theory for a modal language, it is reasonably clear how to obtain a semantic theory from it (Hazen, 1979; Stalnaker, 1986, 1994, dicuss some of the issues). The result departs from standard possible-worlds style modal semantics in four ways:

(i) Distinct worlds have disjoint domains: nothing exists in more than one world.[11]

(ii) There is a collection of binary counterpart relations R that hold only between individuals in distinct worlds.

(iii) Counterpart relations are similarity relations.

(iv) The satisfaction of modal formulae by objects is (a) relativised to counterpart relations and (b) determined by the satisfaction of the corresponding non-modal formulae by their counterparts:

> x satisfies $\ulcorner \Box A \urcorner$ relative to counterpart relation R iff everything x bears R to (its R-counterparts) satisfies A.

[11] At least, no particulars wholly exist in more than one world: multiply located universals, if there are such things, would be wholly present in multiple worlds.

Closed sentences are evaluated for truth by selection of an appropriate counterpart relation.

We want to interpret modal talk counterpart-theoretically to make sentences like 'something is not a necessary part of Tim' true, despite 'Tim' referring to a lump of matter. This allows us to drop theses (i) and (iii) of Lewisian counterpart-theory (though we are not compelled to), modifying (ii) and (iv) thus:

(ii′) There is a collection of four-place counterpart relations R: x in world w, is an R-counterpart of y in world w'.[12]

(iv′) x satisfies $\ulcorner \Box A \urcorner$ at world w relative to counterpart relation R iff, for any object y and world w', if y in w' is an R-counterpart of x in w, then y satisfies A in w'.[13]

On this view it can be true that some of Tim's parts are not necessary parts, despite 'Tim' referring to a lump l of matter. The reason is that l's Tim-counterparts not have the same parts as it(s lump-counterparts): that x in w is an R-counterpart of y in w', does not imply that $x = y$.

Replacing worlds with times gives a temporal version of counterpart-theory. On this view, it can be true that Tim has some non-permanent parts, despite 'Tim' referring to a lump l of matter. The reason is that l's past and future Tim-counterparts need not have the same parts as it(s lump-counterparts): that x at time t, is a temporal R-counterpart of y at t', does not imply that $x = y$.

Say that an object *persists* iff it exists at more than one time. One form of temporal-parts theory, namely *perdurance-theory*, identifies ordinary persistents with fusions of momentary objects. Another version, *stage-theory*, identifies ordinary persistents with momentary objects themselves (Sider, 2001a; Hawley, 2001). In order to make ordinary *de re* temporal discourse true, stage-theory needs (some variant on) temporal counterpart-theory. It should be noted however that temporal counterpart-theory does not require temporal-parts theory. Without temporal-parts, the view is akin to Roderick Chisholm's (1976, ch.3) theory of *entia successiva*:

[12] Counterpart-relations are four-place because (a) lumps exist in more than one world, and (b) a single lump might be Tim in one world, Tom in another, and Tim and Tom's modal properties differ.

[13] This needs complicating to permit non-trivial iterated modality, but it will do as a first pass.

the relation by which we trace the path of an ordinary object through time is not *identity*, but a continuity relation connecting different (mereologically invariant) objects at different times. (Chisholm also endorses the additional thesis that ordinary mereologically variable persistents are logical fictions.)

Given counterpart-theory, differences in the truth-values of modal and temporal claims about lumps and clouds can be attributed to differences in the counterpart relations relative to which those claims are evaluated, rather than the subjects of those claims. To the counterpart-theorist, differences in the truth-values of modal claims about clouds and lumps therefore don't show lumps are not clouds, or that lumps and clouds don't closely resemble one another in cloud-respects. But why should we be counterpart-theorists? What are the motivations for counterpart-theory? Such radical departure from standard modal and temporal semantics is ill-motivated, if its only purpose is to allow retention of a matter-only ontology.

Sider (2001a, ch.5) argues that stage-theory, and hence counterpart-theory, provides the best unified response to the so-called "paradoxes of coincidence": apparent cases in which several objects fill and fit within the same region at the same time (or even throughout time). But it is far from clear why coincidence is supposed to be problematic. Lewis (1986b) has a different motivation for modal counterpart-theory. He is forced into it by his ontology of concrete possible worlds. He also argues that it is a component of the most satisfactory solution to a wide range of puzzles. But Lewis's ontology of concrete possible worlds controversial, to say the least. And there are alternative responses to all of the puzzles Lewis addresses. A counterpart-theoretic defence of Unger's puzzle will be of very limited interest, if these are the motivations for counterpart-theory. I do not claim that these are the only motivations for counterpart-theory. They are however, some of the best and most prominent.

Setting worries about its motivation aside, all forms of counterpart-theory face similar objections. Firstly, counterpart-theory implies that our ordinary judgements of cross-time or cross-possibility sameness are not judgements of identity. Secondly, and relatedly, utterances of *de re* predications at different times have different subjects: the present truth-condition of 'Nick is typing' is that a certain object is typing, while five minutes hence it will be that some other object is typing.

Thirdly, and most significantly, it is doubtful whether modal counterpart-theory is consistent (Stalnaker, 1986, §2). The present version denies that there are possibilities where you have any parts other than your actual parts, whilst also maintaining that you could have had different parts than you actually do. Without an account of possibilities other than as *possible* ways things *could* be, or similar, this is contradictory. The problem concerns the appearance of modal vocabulary in the analysis of possibilities. Counterpart-theory thus brings commitment to a non-modal reduction of modality. It is rightly controversial whether any such reduction is possible. Temporal counterpart-theory does not obviously suffer this last objection because our access to non-present times may not be mediated by temporal vocabulary in quite the same way as our access to non-actual possibilities is mediated by modal and counterfactual vocabulary: we can *remember* the past, but not mere possibilities. (For scepticism about this difference, see Edgington, 2010, §5).[14]

In light of these difficulties and its controversial motivations, let us set counterpart-theory aside.

Second strategy: liberalism about lumps The problem with taking the candidates D_i in Unger's argument as lumps of matter is that lumps and clouds have different temporal and modal properties. This section presents and rejects a more liberal conception of matter on which this is not the case.

Consider the following (partial) theory of matter:

> For every filled spatiotemporal region r, some lump occupies exactly the points in r.

For each cloud, this principle delivers a lump of matter that occupies exactly the same points as it throughout the cloud's history. Since this lump and cloud have the same temporal profile, that provides no bar to identifying them. The objection to the use of PMD1 in Unger's argument therefore fails.

[14] Even if counterpart-theory is defensible, Fine (2003) highlights significant non-modal and non-temporal differences between ordinary objects and lumps: a statue, unlike its constituent lump, may be Romanesque; a cat, unlike its constituent lump, may purr. Counterpart-theory does not address these cases.

This is not our ordinary conception of matter. Suppose that a certain statue has exactly the same parts throughout its existence. Does a lump come into existence when the statue does, occupy the same space as the statue throughout the first half of the statue's history and then cease to exist? Insofar as our ordinary conception of matter speaks to this question, the answer seems to be negative. Yet the principle above entails a positive answer. That principle does not govern our intuitive conception of matter and can therefore provide no intuitive support for the existence of the candidates in Unger's argument. Defending Unger's puzzle by appeal to controversial theses about matter limits its interest; it becomes a puzzle for certain theories of objects, not for our ordinary world-view.

There is a second problem for lumps governed by the principle above: it is silent about their modal profiles. If lumps have only necessary parts, then lumps and clouds will still differ significantly in modal respects relevant to their being clouds. But if we allow lumps to have some merely contingent parts, then the sense in which we have really reduced objects to lumps of matter is unclear; this looks more like a reduction of matter to objects (in combination with a plenitudinous view of objects). The plausibility of Unrestricted Fusion for matter can then lend no intuitive support to the existence of the candidate D_i's.

A response to both these worries is available. Our ordinary conception of matter is inegalitarian: every way of carving up a filled region into subregions is equally legitimate, in the sense that, for each such sub-region r, some lump fills and fits within r. We can respond to the worries above by taking this inegalitarianism as an analysis of our ordinary conception of matter and combining it with a liberal view of regions, to give:

> For every function f from worlds w onto filled spatiotemporal regions of w, there is a lump that, in w, occupies all and only the points in $f(w)$.[15,16]

This provides provides, for each cloud c, a lump of matter that is necessarily co-located with c. Differences in modal and temporal properties provide no bar to

[15] Treat regions as sets of points. When $f(w) = \emptyset$, f defines an object without spatiotemporal location in w. Whether this object exists in w depends on whether spatiotemporal location is necessary for being something.

[16] The Plenitude Lover in Hawthorne (2006c) endorses this thesis.

the identification of such lumps with their coincident clouds. The objection to Unger's use of PMD1 then fails. However, this response still rests on a controversial analysis of matter. A version of Unger's puzzle is of limited interest if it relies on this analysis to defend the existence of the candidates D_i.

A slightly different worry afflicts both strategies. Matter is extensional: whether there is at least one lump for every filled spatial region, or filled spatiotemporal region, or function from worlds onto filled spatiotemporal regions, there is no more than one. Consider the original, and simplest, account of lumps, and the view that identifies ordinary objects with them. Let n be the lump that is now Nick. Then the object-language argument from 'Nick used to be made of different matter' and 'it is not the case that n used to be made of different matter' to 'Nick $\neq n$' must be invalid. Kit Fine (2003) argues that this has untenable consequences in the philosophy of language. And Fine (2000) presents an example of two necessarily co-located objects of which different things are (apparently) true. On each theory of lumps canvassed here, the corresponding argument from these predicative differences to the distinctness of these objects must be invalid. Each of these theories of lumps therefore incurs these untenable consequences, if Fine's arguments are sound. This provides reason to be sceptical of any reduction of objects to matter, and hence also sceptical of any defence of Unger's use of PMD1 that appeals to such a reduction to defend the candidates D_i and cloud C are closely alike in cloud-respects. In the absence of a detailed examination of Fine's arguments, this is certainly not conclusive. We won't however undertake that examination here because an alternative defence of Unger's argument is available.

Third strategy: modify the principle of minute differences We've rejected non-standard semantic theories and analyses of matter as ways of defending Unger's argument against the claim that no cloud and lump are closely alike in cloud-respects. This section presents a version of that argument that accepts these differences and employs an alternative principle of minute differences instead.

First we need a new dyadic relation: *constitution*. This is the relation between a lump and whatever it makes up, or constitutes, e.g.: between the matter of your body and your body, or the marble from which a statue was carved and the statue.

Begin by defining two notions of coincidence:

> *x materially coincides with y* iff x and y both fuse some set s.
>
> *x spatially coincides with y* iff x and y both fill and fit within the same region of space.

The following biconditional may well be extensionally correct, whichever form of coincidence it employs:

> x constitutes y iff x coincides with y and x is a lump of matter.[17]

But intuitively, objects coincide *because* they are made out of the same matter. So we should resist taking the right to analyse the left.

Recall our earlier introduction of resemblance in cloud-respects as resemblance w.r.t. having a structure and make-up appropriate to clouds. Similarly, let resemblance in cloud-constituting respects be resemblance w.r.t. having a structure and make-up appropriate to constituting a cloud. Here is a second principle of minute differences:

PMD2 If (i) some typical cloud x is a fusion of a set s, (ii) x is constituted by y, and (iii) some fusion f of a set s' differs only minutely from y in cloud-constituting respects, then: some fusion of s' is a typical cloud (constituted by f).

Now, (i) C is a fusion of s_0, (ii) C is constituted by the lump D_0, which is a fusion of s_0, and (iii) some fusion of each s_i, namely the lump D_i, differs only minutely from D_0 in cloud-constituting respects. So by PMD2: some fusion of each s_i is a cloud (and constituted by lump D_i). The objection to the argument from PMD1 fails because we're now comparing lumps with lump in respects relevant to their constituting clouds, rather than comparing lumps with clouds in respects relevant to their being clouds. Granting that lumps and clouds are not alike in cloud-respects, PMD2 implies only that each set s_i has some fusion that is a cloud, not that this cloud is the lump D_i.

To avoid trivialising this principle, and thereby rendering it dialectically ineffective, resemblance in cloud-constituting respects must be restricted to exclude,

[17] How could this be false? A lump would have to coincide with an object it didn't constitute (or that wasn't made from it). This does not appear to be a genuine possibility.

for example, resemblance w.r.t. *constituting a cloud*. It is hard to state the restriction precisely. A restriction to microphysical properties and relations might suffice. But what exactly are microphysical properties?[18] Still, it seems that some such restriction is possible. So let us henceforth simply assume that this is so. (§1.1.4 presents a version of the puzzle that does not require this assumption; see especially §1.1.4.4.)

E.J. Lowe and Mark Johnston both appeal to differences in category between lumps and objects in response to Unger's puzzle: no fusion D_i is a cloud because each is a mere lump and no cloud is a lump (Lowe 1982a,c,b, 1995; Johnston 1992). But which lump constitutes C? Lowe and Johnston reply that constitution is vague and flesh this out along broadly supervaluationist lines. We return to supervaluations in chapter 2 and constitutional vagueness in chapter 4 (and Lowe and Johnston's proposal in §4.1). In the meantime, it suffices to note two reasons why the cloud/lump distinction alone cannot solve the problem. (i) It provides no reason to doubt PMD2, and hence no reason to doubt that each D_i constitutes a cloud. (ii) It does nothing to show how, given that each D_i does constitute a cloud, they could all constitute the same cloud.

We now have in place a version of Unger's argument that takes the candidate D_i's as lumps of matter, and uses the principle PMD2 to conclude that each of them constitutes a cloud. This version avoids our objection to the argument that uses PMD1 because it does not assume that any cloud and lump are alike in cloud-respects, or that any lump is a cloud. It also avoids appeal to non-standard semantic theories or analyses of matter. This argument does however, assume that each candidate D_i exists; it doesn't, then PMD2 doesn't imply that it constitutes a cloud. That these candidates do exist follows from Unrestricted Fusion for matter, but although that principle is plausible, it is certainly not beyond reproach. The next section presents a version of Unger's argument that dispenses with the individual D_i's altogether. This shows that Unger's puzzle is not a puzzle about fusion,

[18] Merricks (2003, ch.2, §§II–IV) takes an argument akin to the Problem of the Many to refute the supervenience of mental properties (and composition) on the intrinsic microphysical properties and relations of collections of atoms. We make no assumptions about intrinsicality. So Merricks's argument is silent about PMD2.

constitution or the existence of individuals, but about the instantiation of ordinary sortal properties.

1.1.2.2 Fusions dispensed with

This section presents a version of Unger's argument that does not assume the existence of any controversial entities. (We won't even assume that any of the sets s_i has a fusion.) So, what might we replace lumps of matter with, whose existence is uncontroversial? The most obvious candidates are sets (though they're existence isn't quite uncontroversial). We will need a new principle of minute differences:

PMD3 If some fusion of a set s is a typical cloud and some set s' differs only minutely from s in cloud-respects, then some fusion of s' is a cloud.

We need to understand resemblance in cloud-respects here as resemblance amongst sets w.r.t. those properties of sets that (non-trivially) determine whether they are fused by a cloud. Orthodoxy implies that these will be extrinsic properties because orthodoxy takes sets to lack spatiotemporal location.[19] Intrinsic change requires spatiotemporal location. So if resemblance in cloud-respects between sets were intrinsic, then if some cloud fused a set s at some time, then, at every time, some cloud would fuse s.[20] This is clearly not so.

Resemblance amongst sets in cloud-respects is extrinsic, and therefore does not concern how those sets are in themselves. It obtains because of some other resemblances that obtain between some other entities. Which resemblances, and which entities? The natural answer is:

> s resembles s' in cloud-respects to the degree that the members of s resemble those of s' w.r.t. their making up a cloud.

Two questions arise. Firstly, what is this talk of "the members of s"? And secondly, what is it for those members to make up a cloud?

[19] A property if extrinsic iff it is not intrinsic. A property is intrinsic iff it concerns how an object really is, "considered in itself". The intended contrast is between the extrinsic *being an uncle* and the intrinsic *being a man*. The proper analysis of intrinsicality is a vexed issue we will not enter into here.

[20] The problem is even worse if sets are necessary existents.

The obvious answer to the first question is that this is plural talk: 'the members of s' denotes (plurally) every member of s and nothing else. Following George Boolos (1984), let us take plural expressions to denote several members of the domain from which the denotations of singular expressions are drawn. Plural expressions do not denote plural individuals, but plurally denote many individuals.[21]

With this in place, we can answer our second question above by extending fusion to plurals:

> x is a *fusion of* the y's (also: x *fuses* the y's; the y's *compose* x) iff (i) everything that overlaps x overlaps (at least) one of the y's, and (ii) everything that overlaps any of the y's overlaps x.

Some things make up a cloud iff it is a fusion of them. Note also that the second argument-place of plural fusion is collective:

> F is *distributive* iff, necessarily, if the y's are F, then each y is F.

> F is *collective* iff F is not distributive.

Other collective properties include *being arranged in a circle* and *carrying the boat*. Given a resemblance relation amongst pluralities that is collective in both argument places, our initial gloss on resemblance amongst sets in cloud-respects becomes:

> *s resembles s' in cloud-respects (to degree d)* iff the members of s resemble the members of s' in respects relevant to their composing clouds (to degree d).

This relation is extrinsic because it holds amongst sets in virtue of the (intrinsic) properties and relations of their members, not those of the sets themselves (and sets can be intrinsically invariant despite intrinsic variation in their members).

The appeal to sets here is clearly redundant. Let resemblance in cloud-respects amongst collections be resemblance amongst collections in respects relevant to their composing clouds. We can now state a fourth principle of minute differences:

[21] Subsequent talk of collections and pluralities should be understood as grammatically singular but semantically plural talk about the objects amongst those pluralities and collections.

PMD4 If the x's compose a typical cloud and the y's differ only minutely in cloud-respects from the x's, then the y's also compose a cloud.

This is no less plausible than any of PMD1–3. Let the s_i's be the members of set s_i. For each i, the s_i's resemble the s_0's extremely closely in cloud-respects. Since the s_0's compose our typical cloud C, PMD4 implies that the s_i's compose a cloud D_i. When $i \neq j$: $D_i \neq D_j$ because each overlaps a water droplet disjoint from any the other overlaps. So there are millions of clouds where we thought there to be just one, each nearly coincident with our original cloud C.

The dispensability of assumptions about fusion shows that the Problem of the Many cannot be solved simply by denying the existence of various individuals or restricting fusion. In fact, we can show that it is not even primarily a problem about the existence of ordinary individuals at all, but about the ordinary kinds or sorts to which those objects belong. Suppose the following is true:

Compositional Nihilism Only singletons have fusions (and then only in the trivial sense that every object x overlaps exactly those things that overlap x, including x itself).

It does not follow that clouds do not exist, only that if they do, then either (i) they have no proper-parts, or (ii) *cloud* is collectively instantiated by the pluralities of objects that we would ordinarily say compose clouds. Consider (ii) and the following principle of minute differences:

PMD5 If the x's are collectively a typical cloud and the y's differ only minutely in cloud-respects from the x's, then the y's are also collectively a cloud.

A version of Unger's argument assumes there is just one (typical) instantiation of *cloud* in the sky (by the members of s_0), and uses PMD5 to conclude that there are many such (one by the members of each s_i). The problem therefore remains, despite the fact that no object is a cloud. Since Compositional Nihilism does not solve the problem, it is not a problem about the existence of composite objects.

Since the existence of the fusions D_i is inessential to Unger's argument, it will do no harm to speak as if they do exist, or as if they were candidates to be clouds (rather than merely constitute clouds), in the remainder. Our discussion can always be reformulated in terms of plurals and PMD4, in place of fusions and PMD1.

1.1.3 Principles of minute differences

We've got a version of Unger's argument in place that doesn't rely on any controversial assumptions about the existence of objects like the D_i's. The arguments key premiss is the principle of minute differences PMD4. This section addresses following question of whether principles like PMD1–5 are true. We use 'PMD' as a generic term for all such principles.

We begin by rejecting two arguments against PMD, before turning to two positive arguments for them. Although these arguments aren't decisive, they do reveal that there's much work to be done before rejection of PMD can provide a satisfactory response to Unger. The next section then presents a variant on Unger's argument that doesn't require PMD.

1.1.3.1 Two bad arguments against the principle of minute differences

This section dispenses with two bad arguments against PMD. The first is that PMD is a tolerance principle, and hence known to be false. A tolerance principle for G has the form:[22]

> If x and y differ minutely w.r.t. F and x is a G, then y is a G.

A PMD for G is not of this form, but rather:

> If x and y differ minutely w.r.t. F and x is a typical G, then y is a G.

Say that individuals $x_1, \ldots, x_n$ are a *Sorites series* for G iff (i) x_1 is a paradigm G, (ii) x_n is a paradigm non-G, and (iii) each x_i differs only minutely from x_{i-1} in respects relevant to G (where $1 < i \leq n$). Since F in a tolerance principle is a respect relevant to G, a Sorites series for G is a counterexample to a tolerance principle for G. But a Sorites series for G is not a counterexample to a PMD for G; for a PMD for G implies that x_2 is a G, not that it is a typical-G, and hence implies nothing about x_3–x_n. The (possible) existence of a Sorties series for G therefore refutes a tolerance principle for G without refuting the relevant PMD. Since an appropriate Sorites series can be constructed for most, if not all, ordinary sortals, those sortals are not

[22] Our tolerance principles are a variant on those introduced by Wright (1976).

governed by tolerance principles. It remains an open question whether they are governed by PMD.

We now address the second bad argument against PMD. This argument claims that there are counterexamples to PMD that are independent of Unger's puzzle. Consider the set whose members are our typical cloud C's constituent water droplets and one atom of the British Museum. Hud Hudson (2001, p.26) observes that the fusion of this set resembles C very closely, but is clearly not a cloud. Brian Weatherson (2009, §7.2) cites this as a counterexample to PMD .[23] But this only shows that a more nuanced understanding of cloud-respects is required. Two objects may be very similar overall, despite being highly dissimilar in some more specific respect. The factors relevant to being a cloud, and hence to similarity in cloud-respects, are weighted: large spatial discontinuities count strongly against being a cloud (resembling in cloud-respects), despite counting for little overall difference.

1.1.3.2 Two justifications for the principle of minute differences

Should we endorse PMD? This section considers two reasons to do so. Although neither is decisive, they do revel that rejection of PMD is not an easy response to Unger's challenge.

The first reason to endorse PMD is that it seems analytic. Ignoring Unger's puzzle, PMD seems beyond reproach. Indeed, it is not merely attractive, but plausibly partially constitutive of being a *typical* cloud: how could something that differs minutely in relevant respects from non-cases be a typical, or paradigm, case? Of course, the principle's falsity is compatible with its being intuitively compelling. Ignoring Russell's paradox, the naïve comprehension principle is compelling:

> For any (possibly complex) predicate F in the language of set-theory, $\{x : Fx\}$ exists.

But Russell's paradox refutes this nonetheless. Maybe Unger's puzzle refutes PMD. Still, a solution that retained it would be, *ceteris paribus*, preferable to one that did

[23] Weatherson's purpose is not Hudson's. Hudson uses this case to illustrate how something's being a cloud is sensitive to otherwise small differences, and hence the sensitivity of cloud-respects to those differences, whereas Weatherson thinks that small differences in cloud-respects can ground large differences w.r.t. being a cloud.

not. Furthermore, PMD encodes a conception of paradigms at least as compelling as naïve comprehension for sets.[24] If Russell's paradox is really a paradox, then so is the Problem of the Many, if giving up PMD is what it requires.

Unger (1980, p.161) suggests a second reason to grant PMD: it follows from the vagueness of 'cloud'. According to Unger (1979, §2; 2006a, appendix to ch.7), each vague concept G obeys a *vagueness condition*:

> For some dimension, or respect, F, sufficiently small differences w.r.t. F cannot differentiate a G from a non-G.

This is a tolerance principle by another name. Although he is not explicit about just how this leads to PMD, the idea seems to be this. Were PMD false, then some typical cloud would differ minutely in cloud-respects from a non-cloud. This is incompatible with tolerance for clouds, and hence with their vagueness. So clouds obey PMD.

This is dubious. We've already seen that if a Sorites series for G is possible, then the relevant tolerance principle for G is not a conceptual truth. Since such a series is typically possible, vague concepts are either incoherent or do not obey tolerance principles.[25] Unger endorses the first disjunct. Since most, if not all, ordinary concepts are vague, his view is both problematic and highly controversial (for discussion, see Williamson, 1994, ch.6). So let us set it aside. We should reject tolerance principles, and with them this second justification for PMD. But this is no easy way out. Vagueness is paradoxical precisely because tolerance principles are compelling. Granted that tolerance principles imply PMD, the Problem of the Many is no less problematic than the Sorites, if rejection of PMD is what it requires.

Both justifications for PMD rest on intuitively compelling claims. (i) PMD is analytic of 'typical cloud'. (ii) PMD follows from the tolerance principle underlying the vagueness of 'cloud'. The lesson of Unger's puzzle might well be that these otherwise attractive theses about typicality and vagueness are false. This is not to say that solving the puzzle will be easy. Accounts of vagueness and typicality

[24] Indeed, PMD probably enjoys considerably stronger intuitive support than does naïve comprehension, because set-theory enjoys little or no intuitive support.

[25] There is a third option: the logic of vagueness is weaker than classical, and even intuitionistic, logic. We won't consider this radical view.

that violate (i) and (ii) are required, alongside an explanation for why these false claims seem compelling. However, even rejecting PMD cannot solve the problem. §10 of Unger's original article presents an alternative route from many candidates to many clouds, and the hundred-page discussion in Unger (2006a, ch.7) does not mention PMD at all. To this we now turn.

1.1.4 Selection and exclusion principles

In §10 of "The problem of the many", Unger presents a variant on his original argument that doesn't rely on an appeal to PMD. This section presents this variant, followed by three kinds of inadequate response.

1.1.4.1 The alternative argument

Recall the lumps D_i that are fusions of the sets s_i of water droplets in the vicinity of our cloud C. If C is the only cloud in the sky, then exactly one D_i constitutes a cloud. Which? The alternative version of Unger's argument is driven by two difficulties concerning this question. The first concerns answering it. The second concerns seeing how there could even be an answer.

A *selection principle* provides a property possessed by exactly one D_i, and in virtue of which it constitutes a cloud. An *exclusion principle* provides a property possessed by all bar one D_i, and in virtue of which they don't constitute clouds. Unless there are such principles, either all or none of the D_i's constitute clouds; they differ too little in relevant respects for only one to constitute a cloud. In order for only one candidate to constitute a cloud, a selection principle is required to privilege it over all others, or an exclusion principle to rule out all candidates other than it. Unger claims that there are no such principles, and hence that each D_i constitutes a cloud.

If C is the only cloud, then one true selection principle provides the property of *constituting a cloud*. And true exclusion principles provide the properties of *not constituting a cloud*, and *substantially but not totally overlapping a cloud*. But these are obviously either trivial or circular.[26] Are there non-trivial and non-circular

[26] Circular, in the sense that if x fails to constitute a cloud only because x substantially but not

alternatives?

Because of how closely the D_i's resemble one another, the only candidates seem to involve either the identities of particular candidates, or very fine-grained descriptions of their microphysical make-up.[27] The former fails because any lump that actually constitutes a cloud could have failed to do so. The latter fails because (i) not all clouds are microphysical duplicates of C, and (ii) some D_i that doesn't constitute a cloud would have done so had the D_j that actually does so not existed (because its "extra" droplet hadn't existed). A selection principle that accommodated (i) and (ii) would have to be of the form: in conditions C_1, lumps with property F_1 constitute clouds; in conditions C_2, lumps with property F_2 constitute clouds.... This is problematic because (a) the constitution of clouds should turn on general features instantiable in a range of circumstances, and (b) it is hard to believe that cloud-constitution turns on microphysical structural properties so fine-grained as to distinguish some D_i from all others. Surely reality does not contain substantial distinctions grounded in such slim differences.

We now consider two potential non-trivial and non-circular alternatives: a selection principle in §1.1.4.2, and an exclusion principle in §1.1.4.3. I know of no alternatives. So §1.1.4.4 examines a position that rejects Unger's demand for selection and exclusion principles.

1.1.4.2 Maximality

Sider (2001b, 2003) notes that ordinary sortal properties like *cloud* are maximal:

> "A property, F, is *maximal* iff, roughly, large parts of an F are not themselves Fs."(Sider, 2001b, p.357)

Large parts of houses are not themselves houses, large parts of people are not themselves people, and large parts of clouds are not themselves clouds. Can this solve Unger's puzzle? It seems not.

totally overlaps a cloud y, the question is only pushed back to: why is it y, rather than x that's disqualified from constituting a cloud? Further selection or exclusion principles are then required.

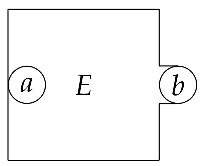
[27] These principles will have to be so fine-grained as to distinguish between lumps that differ by only a pair of water droplets close to their boundaries.

Not all ordinary kinds are maximal. Richard Sharvy (1980) considers a table made by putting two smaller tables together; the smaller tables do not go out of existence or cease to be tables. Popes have worn crowns (the Papal Tiara) comprising two or even three distinct crowns (Wiggins, 1980, p.73). There are two reasons, however, why we should be reluctant to reject maximality on the basis of such examples. The first is that it is hard to find counterexamples involving non-artefactual kinds, and it is these for which Unger's puzzle is most pressing. The second is that these are not obviously counterexamples. A counterexample to maximality for F is a situation containing $a_1, \ldots, a_n$, all of which are Fs, and where (i) $a_1, \ldots, a_{n-1}$ are large parts of a_n, and (ii) the correct answer to the question "How many Fs?" is "n". Sharvy's table and the Papal Tiara plausibly fail condition (ii). Asked how many tables there are, Sharvy could answer "One", or "Two", but not "Three" (and certainly not without qualification).[28]

Recall our original candidates C and D:

Although they almost entirely overlap, neither candidate includes the other. So maximality disqualifies neither from constituting a cloud, and hence doesn't solve the problem.

Maximality will, however, reduce the extent of the problem. Let E be a fusion of D and E:



If E is a candidate, then maximality excludes both C and D. More generally, if there is a unique largest candidate—a unique candidate that includes all candidates that include it—then maximality ensures that only it constitutes a cloud. But there is no guarantee that there will be a *unique* largest candidate. There are two ways to

[28] Rumfitt (2002) contains relevant discussion.

see this.

Firstly, why think that E is a candidate? The fusion of two candidates is not generally a candidate. What difference does the extent of their overlap make? Why believe that the fusion of near-coincident candidates will always be a candidate? Without reason to do so, we lack reason to believe that maximality will narrow the candidates down to one.

Secondly, suppose E is a largest candidate. Some object F differs from E by (i) including some droplet just in E's exterior, and (ii) excluding some droplet on E's boundary. Neither E nor F includes the other. Just the same reasoning that led us to recognise D as a candidate given that C is a candidate, should lead us to recognise F as a candidate given that E is a candidate; for supposing E to be a largest candidate is silent about the underlying problem, namely that many nearby objects are extremely similar in all relevant respects to whichever object possesses whichever property concerns us.

The point is that, whatever it takes to be a cloud, Unger's puzzle already concerned it: many objects in the vicinity of each typical cloud resemble it so closely that it seems arbitrary for just one of them to be (constitute) a cloud. Identifying the cloud with a fusion of what were previously regarded as cloud-candidates does not undermine this, but merely changes the topic, diverting attention to a new puzzle.

Even setting aside these concerns about whether maximality reduces the candidates to one, it is unclear whether it would solve the problem by doing so. The issue turns on whether we understand maximality semantically, or metaphysically.

On a semantic construal of maximality, it governs the application of predicates and concepts: nothing satisfies 'cloud' if it is part of something else that does so, even if it is otherwise just like something that satisfies 'cloud'. On this reading, maximality implies extrinsicality.

The metaphysical construal governs the sortal property *cloud*: nothing that instantiates *cloud* is part of something else that does so. This is neutral about intrinsicality. The property *cloud* may be maximal because whether an object instantiates it depends on the object's external environment. In that case, *cloud* will be extrinsic. But another way in which *cloud* could be maximal is for the boundaries of its possessors to vary depending on their external environment; in which case *cloud* may be intrinsic.[29]

If Unger's abundance of near-coincident clouds gives rise to any genuinely metaphysical problems—i.e. to problems that wouldn't have arisen had we used words differently and that can't be resolved by a more nuanced conception of word-world relations—then semantic-maximality cannot help, even by reducing the candidates to one. (For related discussion, see §3.3.1). So far, the only problem concerns radical error in ordinary numerical judgements. Other problems will be presented in §1.4. If these are genuinely metaphysical, then no purely semantic techniques can resolve them. Unlike semantic-maximality, metaphysical-maximality can help with such problems; but we should doubt whether even that will reduce the candidates to one.

Maximality is guaranteed neither to reduce the candidates to one, nor to solve the problem even if it did so. So let us set it aside.

1.1.4.3 Massive overlap

Clouds plausibly satisfy:

> If x and y massively overlap, then they are not both clouds.

Like maximality, this exclusion principle can be understood either semantically or metaphysically; the last section's discussion carries over wholesale.

Most ordinary kinds seem to satisfy this exclusion principle. It implies that at most one candidate constitutes a cloud. But it provides no reason to think that any particular one does, in preference to all others. Unless supplemented with a uniquely satisfied selection principle, it therefore implies that no candidate constitutes a cloud. Massive overlap alone cannot solve the problem.

1.1.4.4 Brutalism

We've just seen an unsatisfactory selection principle (maximality) and an unsatisfactory exclusion principle (massive overlap). It is not clear what other candidates might be appealed to. So this section examines a view that rejects Unger's demand

[29] Sider (2001b, §1) claims that one adequacy constraint on analyses of intrinsicality is that maximality implies extrinsicality. He does so because he is assuming a semantic conception of maximality.

for selection and exclusion principles. In doing so, we'll clarify Unger's argument and just what that demand amounts to.

Consider (a version of) Peter van Inwagen's (1990) Special Composition Question (SCQ):

SCQ Under what conditions does a set have a fusion?

Ned Markosian (1998) offers the following answer:

Compositional Brutalism There is no true, non-trivial and finitely long answer to SCQ.

Markosian claims that compositional facts are, in this sense, brute facts. Granted this, he claims, Unger's puzzle has an easy "solution": exactly one set s_i of droplets in the vicinity of C has a fusion, this fusion is the only cloud in the sky, and there is no finitely statable non-trivial reason why this set was selected and all others excluded.

Compositional Brutalism is not what's doing the real work here. We saw that Unger's puzzle is primarily a puzzle about the instantiation of ordinary kinds, and only indirectly about composition. To get a puzzle about composition, we need to assume first that ordinary objects are composite individuals (since the problem arises even if Compositional Nihilism is true), and second that all reasonably large composite objects belong to ordinary kinds (since the problem arises even if Unrestricted Fusion is true).

To see what's really doing the work, consider this K-Constitution Question (KCQ):

KCQ Under what conditions does a set have a fusion that belongs to the ordinary kind K?

The analogue to Compositional Brutalism is:

K-**Brutalism** There is no true, non-trivial and finitely long answer to KCQ.

K-Brutalism is compatible with very liberal theses about fusion, but offers a "solution" to Unger's puzzle: exactly one set s_i has a fusion that belongs to the kind *cloud*, though there is no finitely statable non-trivial reason why that set was selected and all others excluded.

One objection to both forms of Brutalism is that the relevant facts are law-governed. Clouds, cats and other ordinary objects do not simply pop in and out of existence randomly, but in a regular and highly systematic manner. Microscopic particles have to be appropriately arranged in order for them to compose (or cease to compose) a cloud. A statement of the laws connecting these arrangements to the existence of clouds would answer those versions of SCQ and KCQ that concern clouds.

The objection fails because the Brutalist can consistently grant that cloud-constitution is law-governed, alongside either of the following theses. (i) The laws are so complex as to resist finite non-trivial statement. This makes our inability to state non-trivial selection and exclusion principles a consequence of our epistemic, cognitive and practical limitations. (ii) Although finitely and non-trivially statable, the laws serve only to delimit a class of candidates; the question of which member of this class constitutes (or composes) a cloud has only a trivial answer. We consider the Brutalist of kind (ii) before returning to (i).

One might object that thesis (ii) makes it arbitrary which s_i composes a cloud. What is the relevant sense of arbitrariness? Suppose that only s_0 composes a cloud, and hence that D_0 either is, or constitutes, a cloud. In what sense is this arbitrary? The Brutalist posits a significant macroscopic difference between D_0 and each other D_i, despite their being no correlated significant microscopic difference. Why should this be objectionable? Why must all significant distinctions be revealed by microphysical descriptions? Relatedly, we can see the Brutalist as opposing the demand for an analysis of macroscopic kinds in microscopic terms. Since we shouldn't expect any such analysis, the Brutalist may claim, we shouldn't expect non-trivial selection and exclusion principles in the first place.

Were the demand for selection and exclusion principles motivated by a demand for a microscopic analysis of macroscopic phenomena, this would be an effective reply. But it need not be so motivated. It is better to see Unger as presenting a challenge to our ordinary world-view: how could only one candidate constitute a cloud, given their extremely close similarity? On what grounds do we retain our belief in only one cloud when presented with the candidates and their close similarities? Plausible selection and exclusion principles would provide the best grounds for doing so. This challenge is not met by simply blocking the argument from the impossibility of stating non-trivial selection and exclusion principle to an abundance of clouds. For we may still draw the disjunctive conclusion that either (a) there are many clouds where we thought there to be one, or (b) there is only one cloud, though there is no non-trivial reason why it is constituted by, say, D_1 rather than D_{560}. An adequate response to Unger's challenge must provide reason to endorse disjunct (b) over (a). Brutalism does not.

This exposes the flaw underlying Brutalism of form (i). The Brutalist is surely right to claim that our inability to find non-trivial selection and exclusion principles *may* be a result of our own limitations. But why should we believe that it is? Brutalism alone provides no reason to do so. So Brutalism fails to address Unger's challenge.

Why endorse the Brutalist solution? Perhaps the best reason is due to W.V.O. Quine (1981b). Quine first endorses three principles: (i) there is at least one cloud in the sky; (ii) distinct clouds don't substantially overlap; (iii) material objects are the material contents of filled spatiotemporal regions. He argues from these to the Brutalist solution, via:

Bivalence For any statement A, either A is true or A is false.

Here's Quine's argument:

> "[W]e are committed... to treating the table as one and not another of this multitude of imperceptibly divergent physical objects. Such is bivalence.... If the term 'table' is to be reconciled with bivalence, we must posit an exact demarcation, exact to the last molecule, even though we cannot specify it. We must hold that there are physical objects, coincident except for one molecule, such that one is a table and the other is not.... In this way simplicity of theory has been served.... [B]ivalence requires us... to view each general term, for example 'table', as true or false of objects even in the absence of what we in our bivalent way are prepared to recognize as objective fact. At this point, if not before,

> the creative element in theory-building may be felt to be getting out of hand, and second thoughts on bivalence may arise." (Quine, 1981b, p.36)

Thus the general methodological principles governing theory-choice, notably the search for simplicity, that motivate Bivalence also motivate the Brutalist solution, though Quine acknowledges that this is a cost of Bivalence.

Now, the Brutalist solution is certainly a cost. Is it one we should pay? Note first that simplicity is only one theoretical virtue amongst many. If this kind of cost-benefit analysis is to motivate the Brualist solution, then the virtues and vices of its rivals must also be assessed. Those rivals include rejection of Quine's (i)–(iii).[30] They also include attempts to mitigate the cost of a non-classical *semantics* by adopting one that retains classical *logic*. Supervaluationism is (sometimes presented as) one such semantic theory. This view is examined in chapter 2, alongside a variant position that retains both classical logic and semantics. Chapter 3 investigates an application of this variant to the Problem of the Many in a manner consonant with Quine's (i)–(iii) and without commitment to Brutalism. General methodological principles can bring commitment to the Brutalist solution only once the alternatives have been properly examined, and hence only upon completion of the investigation we are presently beginning.

1.1.5 Extension to other ordinary kinds

We've focused on clouds. Accepting the puzzle's conclusion—that there are many where we thought there to be one—wouldn't be so bad if it arose only for clouds. This section extends it to all other ordinary kinds of object.

1.1.5.1 Ordinary objects and sorts

Our primary concern in the remainder will be with ordinary objects and the ordinary kinds to which they belong. These objects are the subjects of ordinary talk,

[30] Unger's own response to his puzzle is to reject (i): he takes it so show that our concept of a cloud is incoherent. Accepting an abundance of clouds involves rejecting (ii). Chapter 4 develops a rejection of (iii).

thought and perception, and the kinds to which those subjects belong. Ordinary objects are the most commonplace inhabitants of the everyday world around us, e.g.: grains of sand, bricks, tables and organisms. They also include larger and smaller objects—e.g. viruses, microbes, planets and galaxies—whose recognition requires specialist equipment, provided they have similar internal complexity and coherence to the paradigm cases.

Of all the kinds to which ordinary objects belong, our primary interest is in sortals. Sortal properties are those, like *human*, *cat* and *cloud* that (by and large) suffice for countability and demarcate their bearers from other individuals. For example, although *water* is a natural kind, the question "How many waters are in the cup?", can only be met by mystification and a request for further information: having counted *this* water, should *that* water also be counted? (And how big is *that* water anyways?) Sortals like *cat* and *cloud*, by contrast, automatically supply answers to this question.

It does not follow that "How many F's?" can always be answered when F is a sortal. It may be vague whether x is a cat, or whether x is a different cat from y. It may not even be possible for beings with our limited practical and epistemic capacities to arrive at a settled answer. But these are the exceptions. Ordinary sorts typically permit counts with determinate correct answers. One lesson of Unger's puzzle might be that sortals are more like non-sortals in this respect than it would otherwise seem.

These characterisations of ordinary objects and sortals are both somewhat imprecise. But in combination with the paradigms listed, they will suffice for our purposes. Our concern is not to precisely delimit the notions of sortal and ordinary object, but to show how widespread Unger's puzzle is.

1.1.5.2 The problem for non-clouds

Unger's puzzle arises because almost coincident with each typical cloud are many objects that resemble it extremely closely in cloud-respects. The same applies to other sorts of ordinary objects. They all suffer the same difficulty: what guarantees that exactly one candidate will be better-placed than all others nearby?

In the case of clouds, the problem results from the gradual decrease in droplet density across their boundaries. This ensures that inclusion of a single droplet only just in the cloud's exterior will be both possible and provide something that resembles the cloud extremely closely. For most objects however, mere proximity to their boundaries is insufficient for such close resemblance. The fusion of Tibbles the cat with a dust particle resting on his skin does not closely resemble Tibbles in cat-respects.[31] In order for an entity to be part of a cat, it entity must be appropriately related to other parts of that cat. Unlike clouds, the appropriate relations are not purely spatial, but include causal and attachment relations, and maybe others besides. Whatever the precise details, let us lump these relations together as cat-bonding relations:

> If x is a cat, then y is part of x iff y is cat-bonded to some part of x.

Causal interaction, attachment, proximity and the like are all matters of degree. It is thus very plausible that cat-bonding is also a matter of degree. This is further supported by the observation that cat's hairs don't simply "pop out" between instants, but gradually work their way loose. Similar remarks apply to all other ordinary sorts.

The gradualness of cat-bonding suffices to generate Unger's puzzle. Suppose that hair h has only just ceased to be part of Tibbles; were h any better cat-bonded to Tibbles, it would be part of him.[32] Let Tigger be a fusion of Tibbles with h. Tigger resembles Tibbles extremely closely in cat-respects. A principle of minute differences for cats implies that Tigger is a cat:

> If x is a typical cat and y differs only minutely in cat-respects from x, then y is also a cat.

In addition to this argument via a principle of minute differences, Unger's second version of the puzzle also arises. The prospects of stating non-trivial and non-circular selection and exclusion principles look no better here than in the original case of clouds. So what guarantees that there is a unique most inclusive cat-like

[31] Resemblance in cat-respects is resemblance in those respects relevant to whether something is (or constitutes) a cat.

[32] The problem is exaggerated by considering individual molecules, electrons and the like.

object on Tibbles's mat? Without an answer, our belief that there's only cat on that mat looks no better off than our belief that there's only one cloud in the sky. Since this turns on no peculiarity of cats, the generalisation to all other ordinary sorts is straightforward.

Another way of seeing the problem is as the challenge of matching each macroscopic object with a unique collection of microscopic particles, or infinitely precise region of spacetime. This seems required for Tibbles to be the only cat on his mat; for if there are many equally good ways of associating Tibbles with (not entirely coincident) collections of microscopic particles, then each such collection is equally suited to compose a cat; how, then, could only of them do so? The problem is that our conception of ordinary macroscopic objects doesn't appear to provide for such fine-grained distinctions amongst microscopically individuated collections of particles or lumps of matter: many are ruled out, but more than one remains. Tibbles's boundaries are determined by, for example, the attachment of hairs. But unless the event of Tibbles's hair falling out admits of uniquely correct re-description as an event in which such-and-such microscopic entities (and no others) participated, then the features that determine Tibbles's boundaries will not suffice to distinguish amongst closely resembling lumps (or collections of particles). Until such a uniquely correct microphysical re-description is supplied, there is no reason to believe that one is possible; this is all that Unger's puzzle requires.

The problem is most striking for beings like ourselves. If the argument is sound, then millions of humans are sitting in your chair right now. A similar argument shows that if any of these are persons, then so are all the others: millions of people are sitting in your chair, reading this thesis. Whatever our views about the number of clouds in the sky or cats on Tibbles's mat, this is difficult to take seriously.

1.2 Lewis's puzzle

In Lewis's hands, Unger's puzzle becomes a puzzle of vagueness. Here's his initial presentation:

> "Think of a cloud—just one cloud, and around it clear blue sky. Seen from the ground, the cloud may seem to have a sharp boundary. Not so.

> The cloud is a swarm of water droplets. At the outskirts of the cloud, the density of the droplets falls off. Eventually they are so few and far between that we may hesitate to say that the outlying droplets are still part of the cloud at all; perhaps we might better say only that they are near the cloud. But the transition is gradual. Many surfaces are equally good candidates to be the boundary of the cloud. Therefore many aggregates of droplets, some more inclusive and some less inclusive, (and some inclusive in different ways than others), are equally good candidates to be the cloud. Since they have equal claim, how can we say that the cloud is one of these aggregates rather than another? But if all of them count as clouds, then we have many clouds rather than one. And if none of them count, each one being ruled out because of the competition from the others, then we have no cloud. How is it, then, that we have just one cloud? And yet we do." (Lewis, 1993a, p.164)

Later, he gives another:

> "Cat Tibbles is alone on the mat. Tibbles has hairs $h_1, h_2, \ldots, h_{1000}$. Let c be all of Tibbles including all these hairs; let c_1 be all of Tibbles except for h_1; and similarly for $c_2, \ldots, c_{1000}$. Each of these c's is a cat. So instead of one cat on the mat, Tibbles, we have at least 1001 cats—which is absurd.... Why should we think that each c_n is a cat?... [S]uppose it is spring, and Tibbles is shedding. When a cat sheds, the hairs do not come popping off; they become gradually looser, until finally they are held in place only by the hairs around them. By the end of this gradual process, the loose hairs are no longer parts of the cat. Sometime before the end, they are questionable parts: not definitely still parts of the cat, not definitely not. Suppose each of $h_1, h_2, \ldots, h_{1000}$ is at this questionable stage. Now indeed all of $c_1, c_2, \ldots, c_{1000}$ and also c which includes all the questionable hairs, have equal claim to be a cat, and equal claim to be Tibbles. So now we have 1001 cats. (Indeed, we have many more than that. For instance there is the cat that includes all but the four hairs h_6, h_{408}, h_{882}, and h_{907}.) The paradox of 1001 cats... is another instance

of Unger's problem of the many." (Lewis, 1993a, pp.166–7)

This section examines this argument and its relationship to Unger's puzzle.

With one exception, Lewis's first presentation is close to Unger's. Unger does not claim that all the candidate surfaces have equally good claim to be the boundary of the original cloud. Unlike Lewis, he assumes that the cloud's boundary is settled, and argues for an abundance of equally cloud-like boundaries regardless, though none bounds the original cloud. Why the difference?

The answer lies in Lewis's second presentation: the boundaries differ only by the inclusion and exclusion of "questionable parts", entities neither definitely part nor definitely not part of Tibbles. Because of this appeal to questionable parts, Lewis's conclusion differs slightly from Unger's: although there are many cats on the mat, they all have equally good claim to be Tibbles, they are all questionably Tibbles. What is the sense of "questionable" and "definite" here? To illuminate this, we begin with the phenomenon of vagueness.

1.2.1 Borderline parts

This section examines the notion of a questionable, or borderline, part. We begin with a brief introduction to vagueness generally, before turning to parthood.

1.2.1.1 Vagueness

A theory-neutral characterisation of vagueness is difficult, if not impossible. So we proceed via paradigm cases and identification of some characteristic features.

Vagueness is the "fuzziness" of the distinction between, for example, the red and the orange, the tall and the not tall, the intelligent and the unintelligent, or the chairs and the stools. In a sufficiently well-stocked spectrum from one of these poles to its pair, there is no sharp transition from one to the other, but a fuzzy transitional region. On a colour chart, for example, the red zone does not seem to abruptly terminate when the orange zone begins, but to blur gradually into it.

The cases in this intermediate zone are the borderline cases. When x is a borderline case of F, the appropriate response to the question "Is x (an) F?", seems to be neither a simple "Yes" nor a simple "No". This much is a datum.[33] Exactly what the appropriate response should be is a matter of dispute. We introduce the notion of *clarity* to characterise the borderline cases. The positive cases outside the fuzzy region are clearly cases. The negative cases outside that region are clearly not cases. The borderline cases are neither clearly cases nor clearly not cases. By way of example, brake lights are clearly red, oranges are clearly orange, and terracotta pots are borderline red/orange.

One characteristic of vagueness, as opposed to other forms of unclarity, is that the extent of fuzziness is itself fuzzy. An example: the clearly red zone on a colour chart does not seem to abruptly terminate when the borderline red/orange zone begins, but to blur gradually into it. Fuzziness permeates vague classification. This gives rise to the phenomenon of higher-order vagueness: borderline cases to the borderline and clear cases (and borderline cases to those cases etc.). More on this in §2.9.

Vagueness is also responsible for the Sorites paradox. Consider the claims:

R1 Ten seconds after his birth, Bertrand Russell was young.

R2 Forty years after his birth, Russell was not young.

R3 If Russell was young i seconds after his birth, then he was also young $i+1$ seconds after his birth.

All three are intuitively compelling. Indeed, each is plausibly partially constitutive of the meaning of 'young'. But they are classically inconsistent.[34] For instantiating R3 for $i = 10$ followed by *modus ponens* using R1 leads to the conclusion that Russell was young eleven seconds after his birth. Instantiating R3 for $i = 11$ and another application of *modus ponens* gives the conclusion that Russell was young twelve seconds after his birth. Repeating this process eventually gives the conclusion that Russell was young forty years after his birth, which contradicts R2. This

[33] I do not claim that neither "Yes" nor "No" may be offered in response the question "Is x (an) F?" when x is a borderline F. I claim only that such answers should be qualified, to avoid misleading one's audience.

[34] More precisely: R1–R3 are classically inconsistent given some elementary arithmetic and obvious truths about the structure of time.

is an instance of the Sorites paradox. The following claims provide another:

T1 Anyone two-hundred centimetres in height is tall.

T2 Anyone one-hundred centimetres in height is not tall.

T3 If anyone of i centimetres in height is tall, then anyone of $i-1$ centimetres in height is tall.

Again, repeated instantiations of T3 and applications of *modus ponens* lead from T1 to the conclusion that anyone one-hundred centimetres in height is tall, which contradicts T2. Premisses R3 and T3 express what §1.1.3 called tolerance principles. We will call them *Sorites principles*. Their plausibility results from the vagueness of the property or concept in question, *being young* in our first case, and *being tall* in the second.

The Sorites paradox and borderline cases seem to be connected: the presence of the borderline cases seems to explain the plausibility of Sorites principles. Consider the negation of R3:

$$\neg\forall i(\text{Russell was young } i \text{ seconds after his birth } \rightarrow \text{Russell was young } i+1 \text{ seconds after his birth})$$[35]

This is classically equivalent to:

$$\exists i(\text{Russell was young } i \text{ seconds after his birth } \wedge \neg\text{Russell was young } i+1 \text{ seconds after his birth}) \quad (1)$$

Which, if true, has a true instantiation:

$$\text{Russell was young } n \text{ seconds after his birth } \wedge \neg\text{Russell was young } n+1 \text{ seconds after his birth}$$

What might the cut-off point n be? Since a conjunction entails its conjuncts, we should propose only those answers both of whose conjuncts we are prepared to endorse. There are three cases.

[35] '$\rightarrow$' is the material conditional throughout.

First case: Russell was clearly young n seconds after his birth. Then Russell was either clearly young or borderline young $n + 1$ seconds after his birth. Either way, (unqualified) endorsement of 'Russell was not young $n + 1$ seconds after his birth' is inappropriate. So we should not propose n as the cut-off point.

Second case: Russell was borderline young n seconds after his birth. Then (unqualified) endorsement of 'Russell was young n seconds after his birth' is inappropriate. So we should not propose n as the cut-off point.

Third case: Russell was clearly not young n seconds after his birth. Then endorsement of 'Russell was young n second after his birth' is clearly inappropriate. So we should not propose n as the cut-off point.

In each case, we should not propose n as the cut-off point. So we should not propose n as the cut-off point. Since n was arbitrary: we should not propose, of any i, that i seconds after his birth is the cut-off point for Russell's youth. We've just seen that this follows from the following two factors: (i) outright endorsement, without qualification, of a borderline statement is inappropriate; (ii) the clear positive and negative cases are separated by borderline cases. The explanation for the attraction of R3 is then that we reject (1) because we know *a priori* that we ought not to endorse any of its instantiations; and since we reject (1), we endorse it's negation R3.

This explanation won't do as it stands; for we accept some existential generalisations despite knowing *a priori* that we ought never to endorse any of their instantiations. Two examples: 'some mammal was the first unnamed dog born at sea', and 'something very strange happens inside a black-hole's event-horizon'. Both are relevantly disanalogous to vagueness, and hence don't undermine our explanation for the plausibility of T3. The first turns on semantic vocabulary appearing within the scope of a quantifier, yet none appears in (1). The second turns on our own epistemic limitations, but our ignorance of when Russell ceased to be young does not: it seems, in some sense, misguided even to wonder about, never mind set about trying to discover, when Russell ceased to be young.[36]

[36] Not all will grant this. Epistemicists like Timothy Williamson claim that we are ignorant of when the last second of Russell's youth was (Williamson, 1994). But since not all ignorance results in vagueness, they too must find a disanalogy between typical ignorance and vagueness. The epistemi-

With this preliminary introduction to vagueness complete, we return to Lewis's puzzle.

1.2.1.2 Mereological vagueness

Suppose Tibbles is moulting. Let $h_1, \ldots, h_n$ be a series of hairs where (i) h_1 is clearly part of Tibbles, (ii) h_n is clearly not part of Tibbles, and (iii) each h_i is only very slightly less firmly attached to Tibbles than h_{i-1} (where $1 < i \leq n$). $h_1, \ldots, h_n$ are a Sorites series for being part of Tibbles.

This series seems to exhibit the fuzziness characteristic of vague classification. The hairs that are part of Tibbles don't seem to be immediately succeeded in the series by those that aren't. Rather, some hairs h_m–$h_{m'}$ are borderline parts of Tibbles, separating the clearly attached hairs h_1–h_{m-1} from the clearly detached hairs $h_{m'+1}$–h_n.[37] That there are such borderline hairs is reinforced by the plausibility of the Sorites principle:

> If h_i is part of Tibbles, then so is h_{i+1}.

For if some hairs are borderline parts of Tibbles, then the previous section's explanation for the attraction of Sorites principles generally, can also be used to explain the attraction of the particular principle above.

Each borderline hair provides a borderline cat-candidate (or cat-constituting-candidate). Let T be all of Tibbles, excluding $h_1, \ldots, h_n$; let each T_i amongst $T_1, \ldots, T_n$ be a fusion of $\{T, h_1, \ldots, h_i\}$. When it's borderline whether h_i is part of Tibbles, it's also borderline whether T_i is (constitutes) a cat.

1.2.2 Why many cats?

We now have a range of cat-candidates, namely each fusion T_i of the set $\{T, h_1, \ldots, h_i\}$ where h_i is a borderline part of Tibbles. Since it's borderline whether h_i is part of a cat, it's borderline whether T_i is a cat.[38] This much is unproblematic. The prob-

cist may then use this disanalogy to defend the argument from non-endorsement of each instantiation of (1) to acceptance of T3.

[37] It will, of course, be a vague matter just where in the series h_m and $h_{m'}$ are located

[38] We talk for simplicity as if the fusions could be cats, rather than merely constitute cats. We've seen that nothing turns on this.

lem arises because Lewis draws the stronger conclusion: each candidate T_i is a cat. What licenses this?

Suppose that some candidate is more cat-like than any other. Then the others are surely not borderline cats, but clearly non-cats: the best candidate wins. Since they are all borderline, all are equally cat-like.[39] But one way for one to be more cat-like than any other is for only it to be a cat. Then since at least one of the candidates is a cat, they all must be. So there are many cats on Tibbles's mat.

This argument licenses the following stronger conclusion than Unger's: for each candidate T_i, it's borderline whether T_i is Tibbles, the cat with which we began. Since each candidate has equal claim to be Tibbles, the result is borderline identity sentences 'T_i = Tibbles'.[40] This shows that something must have gone wrong somewhere in Lewis's argument; for an argument exactly parallel from the candidates all being borderline cats to their all being cats can now be used to show that they are all Tibbles. Yet that's impossible because the candidates are many, while Tibbles is one.

What has gone wrong? The answer must be that although the candidates are all on a par w.r.t. being cats and it's clear that one of them is a cat, it doesn't follow that any one of them is clearly a cat; they may all be only borderline cats instead. Let the sentential operator 'Δ' formalise 'It is clearly the case that…'. The following argument-form must be invalid:

$\Delta\exists xFx$, therefore: $\exists x\Delta Fx$.

And the following must be consistent:

$$\Delta\exists xFx \wedge \forall y(\neg\Delta Fy \wedge \neg\Delta\neg Fy)$$

This is a constraint on theories of vagueness if (i) the candidates are all borderline cats, (ii) it's clear that one of them is a cat, and (iii) it's clear that there is only one cat the mat. The supervaluationist views considered later all respect this.

[39] Indeed, that the T_i's are all equally cat-like is what seems to be responsible for their being borderline cases in the first place.

[40] This doesn't conflict with Gareth Evans's (1978) famous argument against vague identity. For an identity sentence can be vague due to vagueness about the referents of the terms flanking '=', rather than vagueness in the identity relation. See Lewis (1988b) for discussion.

1.3 How many puzzles?

Are Unger's and Lewis's puzzles one and the same? It is too early to give a firm answer, but here are some of the issues.

On the assumption that *cat* is maximal, Lewis's puzzle results in more cats than Unger's. Each Lewis-candidate T_i includes each candidate T_j when $i \leq j$. So maximality implies that at most one Lewis-candidate is a cat: when some candidates are otherwise equally suited to be cats and one includes all others, maximality implies that that largest one is the best. So, ignoring vagueness, there is a unique largest best candidate in the series $T_1, \ldots, T_n$. Hence Unger can recognise only this one as a cat. Unger's cats extensively overlap, but don't include one another. Vagueness, however, undermines the thought that there will be a *unique* best candidate in the series of increasingly inclusive candidates: several candidates can have equally good claim to best despite one being clearly largest because it may be borderline whether a smaller candidate is more cat-like than some larger one.

Another difference concerns the conclusions of the puzzles. As we saw in the previous section, Lewis's cats, unlike Unger's, are not clearly distinct from the cat with which we began.

Thus we have two reasons to distinguish the puzzles. Unger's most recent work on the topic argues for a third: his would arise even were there no vagueness (Unger, 2006a, pp.369–70, 468–9, chs.7.8–7.9). Unger claims that there could be many objects on the mat that differ minutely in cat-respects from Tibbles, even were it entirely precise and determinate which of these objects coincides with Tibbles. So Unger's puzzle, unlike Lewis's, does not make essential appeal to vagueness in Tibbles's boundaries.

This argument is only sound if its premiss is true. Is it really possible for Unger's puzzle to arise on the assumption that Tibbles is sharply bounded? Not if Unger's puzzle is a source of vagueness, or if both puzzles are manifestations of a single underlying phenomenon. Thus whether there are two puzzles here or one turns ultimately on the nature of vagueness. It cannot be settled, as Unger wishes, in isolation from an investigation into the nature of vagueness. Having settled on an account of vagueness in chapter 2, chapter 3 examines and rejects an application of it on which Unger's attempt to distinguish the puzzles is unsound; chapter 4 then endorses an application on which Unger's attempt succeeds.

One way to see the issue is this. Unger's is the puzzle of too many candidates: what guarantees that one candidate is better than all others? If there is no such guarantee, how can there be only one cat? Lewis's is the puzzle of borderline candidates: how can there be only one cat when all the candidates are borderline (and hence equally good candidates), and yet one of them clearly is/constitutes a cat? Several questions arise about the relations between these puzzles. Does an overabundance of best candidates imply the existence of borderline candidates? Do the existence of borderline candidates imply an overabundance of best candidates? And supposing that the answer to both questions is "Yes", are these puzzles both manifestations of a single phenomenon, or of distinct yet mutually entailing phenomena? These questions cannot be settled in isolation from an account of vagueness. If it turns out that both puzzles are manifestations of a single phenomenon, then the first two reasons to distinguish the puzzles fail: although Unger may recognise fewer candidates and draw a weaker conclusion than Lewis, that's only because he's failed to recognise the true nature of his puzzle.

1.4 Puzzle or problem?

Why not just accept the conclusion that there are many cats on Tibbles's mat? It conflicts with our ordinary world-view, but that's insufficient for rejection if philosophy can make genuine discoveries. (And what's the point of philosophy if it can't?) This section presents several alternative problems. We do not claim that any is decisive, only that together they make a cumulative case against accepting an abundance of cats on Tibbles's mat.

1.4.1 Time, modality and coincidence

The Problem of the Many is a source of fission and fusion puzzles. Suppose that Tibbles's boundaries are entirely precise at time t, and there are no other nearby cat-candidates; neither Unger's nor Lewis's puzzle arises at t.[41] Suppose also that Tibbles's boundaries are extremely vague by the later time t', so that millions of almost coincident cats are then on the mat. Which later cats were on the mat at t? Tibbles seems to have undergone fission. And if Tibbles's boundaries later become less vague, then those millions of cats seem to undergo fusion. We'll focus primarily on fission, but similar considerations arise in both cases.

Fission cases are often presented as a source of insight about the nature of persistence. They are usually thought to be atypical. But if the reasoning behind the Problem of the Many is sound, then fission is the norm for all kinds of ordinary object, including ourselves. Our view about fission had better not, therefore, conflict with our ordinary judgements of persistence. For example, since cats and persons survive for extended periods of time, this rules out approaches on which fission destroys objects, replacing them with two new individuals; yet that's quite an attractive approach to fission.

Relatedly, fission resulting from the Problem of the Many is unlike "ordinary" fission. Surely cats are not brought into being by hairs working loose from other cats; that's just the wrong kind of change to create a cat. So all the cats on the mat after fission—following an increase in the extent of the Problem of the Many—were on the mat beforehand. We now have an argument not just for the near-coincidence of millions of cats on the mat, but for their (temporary) total coincidence prior to fission.

A modal analogue strengthens this from temporary coincidence to permanent coincidence. Suppose our original fission case occurred in world w. Let v be a world just like w until immediately before Tibbles's boundaries became vague, when Tibbles was destroyed. How many cats are on Tibbles's mat in v? Surely there should be only one; there is no vagueness in the boundaries of any cat on the mat in v. But since cats don't come into existence when hairs fall out of other cats, all the post-fission cats in w must be present in v too, where they coincide throughout their entire lives.

[41] The assumption that Tibbles's boundaries were ever utterly precise is inessential; only the weaker assumption that the extent of their vagueness can vary over time is necessary. The stronger assumption simplifies presentation.

The alternative, that there is only one cat on the mat in v, seems to imply that the number of earlier cats depends on the future course of events. For v is just like an initial segment of w, except only that there are fewer cats in v than in w. Yet surely the number of past cats shouldn't depend on how things turn out in the future.

Something similar applies even if every cat in w is also in v. What reason is there, in v, for the existence of all those cats? There seems to be none, other than to accommodate the (merely) possible future extent of vagueness. Yet surely the number of past and present cats should not depend on how events *could* (and in this case didn't) unfold in the future. This approach will also increase the extent of total coincidence: there are as many cats in v as there are in any possible future that could have unfolded from some time in v. Since Tibbles's boundaries could have, but didn't, become very vague indeed, very many cats are on the mat in v and coincident throughout their entire lives; there is no reason in v for any of these cats to exist other than that events could have (but didn't) unfolded so that they didn't quite coincide with Tibbles.

The near-coincidence of cats in our original version of the problem might not concern us. After all, partially overlapping objects are commonplace. But we have just argued for the permanent total coincidence of Tibbles with millions of cats in v. Even defenders of coincident entities might balk at this. David Wiggins (1968), for example, grants that distinct objects can, and often do, occupy the same place at the same time, but denies that objects of the same kind can do so even momentarily.

Can we avoid these coincident cats in v? Chris Hughes (1986) surveys the options. First option: our description of v is multiply satisfied; really there are millions of worlds qualitatively just like v, each of which contains only one cat, a different one in each. One consequence is that a world's qualitative history plus the identities of everything bar the cat(s) on the mat, is insufficient to determine which cat is on the mat. It is also hard to regard these worlds as genuinely possible. Presumably, any cat would still exist were its boundaries a little more precise. Yet on a standard possible-worlds style semantics for counterfactuals, the current proposal will falsify this.[42]

[42] On the Lewis-Stalnaker semantics: $\ulcorner A \Box\!\!\rightarrow C \urcorner$ is true at w iff every closest world to w at which

Second option: our concept *cat* is not a single concept; millions of different cat-concepts determine just slightly different cat-like paths through modal space. Each cat-concept pairs one object on the mat in w with one on the mat in v, and applies to no more than one such in each. This view replaces our natural kind concept *cat* with many different such concepts. So the candidates don't all belong to the same kind. Yet surely there are not so many natural kinds. And how could these coincident objects that come into existence as a result of the same natural processes and which are capable of breeding with exactly the same objects really belong to different kinds? This abundance of biological kinds is not even discoverable by the standard methods of biologists, but only by *a priori* reflection on the boundaries of cats.

Only this second option is available in the purely temporal version of the problem. There is no space to claim that our descriptions of the past fail to distinguish between an array of qualitatively similar pasts in which different objects exist. So we seem committed to either (a) the dependence of the past on the future and the coincidence of objects of a kind, or (b) implausible differences in kind between the objects on the mat.

1.4.2 Causation

Trenton Merricks (2003) presents the following argument against the existence of most ordinary objects:

(i) If ordinary objects exist, then they cause the same effects as their constituent atoms acting in concert.

(ii) If ordinary objects cause the same effects as their constituent atoms acting in concert, then there is widespread and systematic causal overdetermination.

(iii) There is no widespread and systematic causal overdetermination.

(iv) So there are no ordinary objects.

A is true is a world at which C is true. On the view described in the text, no candidate exists in each closest world to w in which the cat on the mat's boundaries are precise. So the following is untrue in w: 'if Tibbles's boundaries were a little more precise, then any cat on the mat would still have existed'.

Merricks goes on to maintain that only objects whose causal powers go beyond those of their constituents avoid elimination via this argument. He also claims that only conscious objects have this feature, and hence that the only ordinary objects are conscious objects.

This argument is valid. To resist it, we must resist its premisses. Which? Consider premiss (iii). Why deny that there is systematic and widespread causal overdetermination? Though Merricks offers no argument, some will grant it nonetheless. The argument's soundness then turns on premisses (i) and (ii). Yet the Problem of the Many shows that only premiss (iii) is required.

Suppose that many cats almost coincide with Tibbles. Almost anything that any one of these cats causes is also caused by each of the others. So if the argument for many cats is sound, then there is widespread and systematic causal overdetermination. So by premiss (iii) alone: objects susceptible to the Problem of the Many do not exist.

1.4.3 Free will

Hudson (2001, ch.1.5) argues that the Problem of the Many challenges our ability to act freely. Suppose that you are an ordinary material object, a human being say. By the Problem of the Many: there are many humans in your chair. A similar argument shows that since you are a person, so are they: many people are sitting in your chair. Suppose you freely lift your arm. It follows of necessity that each other person in the chair lifts their arm. The following principle then implies that only your action was free:

> "If (i) A's freely doing x at t entails B's doing y at t, and (ii) A freely does x at t, and (iii) A is distinct from B, then B does not freely do y at t." (Hudson, 2001, p.40)

The formulation needs modifying to allow for God's freedom: your freely lifting you arm entails God's letting you do so, but God freely let you do so. We might also question the application of entailment to actions rather than propositions. Still, something along these lines is intuitively very plausible.

It follows that at most one person in your chair acts freely at any time. Which? The reasoning behind the original problem leads to the conclusion that either all or none of them ever acts freely. How could any one be non-arbitrarily selected and all others excluded from acting freely? Any principled selection and exclusion for free action would presumably also suffice to ensure that there is only one person or human in your chair. Since there are many persons in your chair, either all or none of them ever acts freely. By Hudson's principle: at most one does. So none do.

1.4.4 Real choice

Unger (2006a, ch.7) argues that your power to make real choices, choices independent of those of any other person, shows that you are the only person in your chair.

Suppose you have never previously considered either the concept of a blue sphere or the concept of a red cube, and also lack pre-existing inclination towards imagining instances of either concept. (Substitute as required to make this supposition true.) Imagine either a blue sphere or a red cube. Write down which you imagined. Repeat as often as you like. Each time, I assume, you write down whichever you imagined. Unger thinks this counts against the existence of many other people in your chair; for were we to ask millions of people to carry out this experiment, we would expect divergence in their answers. Beings with the power to make genuine choices will tend to make different choices when they lack prior inclination towards one option over another.

Unger consider three responses. First response: we lack the power to make genuine choices; our ability to choose is constrained by the other people we almost coincide with. Second response: it's just pure luck that you and your many always make the same choice, that they never thwart your decisions. Third response: beings with the power to make genuine choices are simple non-physical entities—hence not susceptible to the Problem of the Many—that causally interact with their many human bodies. This is Unger's preferred solution. He misses a fourth response: our power to make genuine choices is constrained by our physical make-up in such a way that near-coincident choosers cannot manifest this power in different ways. On this view, the argument rests on an inadequate metaphysics of choice. This last is perhaps the most attractive option, but each will be objectionable to some.

1.4.5 Moorean fact

Moorean facts are theses whose plausibility is so great that no philosophical argument could refute them. For each Moorean fact M, it is supposedly more plausible that any argument (or collection of arguments) against M involves a false premiss or invalid inference, than that M is false. When we cannot locate this false premiss or invalid inference, it is supposedly more plausible that we are in error than that the argument is sound.

Is it a Moorean fact that, sometimes, only one typical cloud is in the sky? It is certainly very plausible. But belief in the falsity of a Moorean fact is supposed to be so radical, the departure from our ordinary world view so great, that it cannot be seriously entertained for long (or outside the philosophy room). An abundance of clouds where we thought there to be just one does not seem to be of this kind, or to be nearly so radical as the falsity of standard examples, e.g.: there is an external world; I have two hands; $2 + 2 = 4$; murder is wrong.

That Tibbles is the only cat on his mat looks like a better candidate. But is it really impossible to believe otherwise? That you are the only person reading this page looks like a better candidate; it is very strange to think that "you" are not one person but many. But it's still not clear that I can't seriously entertain that thought. On the other hand, there is surely only one person in *my* chair. Nobody else is perceiving my computer screen, considering what I should write next, or wondering what I'll have for lunch. That I am the only person in my chair seems as good a candidate Moorean fact as any. And likewise, *mutatis mutandis*, for you, I suppose.

1.4.6 Responsibility

The Problem of the Many threatens our most commonplace methods of apportioning praise, blame and moral responsibility. Suppose that someone commits murder and is punished with a life sentence. The punishment is out of proportion to the crime: the murderer killed millions of people. Furthermore, if Hudson and Unger's arguments about freewill and choice are sound, then at most one person almost coincident with all those who were punished freely chose to commit the crime. Since one can be justly punished only for what one freely chooses to do, millions are routinely unjustly punished by even the most careful judicial system. Similarly, even the smallest charitable donation can improve the lives of thousands, and monogamy is impossible. It is unclear whether our ordinary moral beliefs and practices can be reconciled with the Problem of the Many for persons.

1.4.7 Singular thought and reference

The Problem of the Many threatens the possibility of singular, or *de re*, thought about particular objects (Unger, 1980, §12A). Suppose Rosie tries to think about a particular book on her desk. No feature of her thought or perception of the book privileges just one of the many with which it almost coincides; nothing about Rosie or the books could disqualify all but one from being the subject of her thought. In what sense is Rosie's thought singular? How can Rosie have a singular thought about a book unless there is some book her thought is about? She surely isn't having a different *de re* thought about each book. (This very last claim is questioned in §3.2.3).

A more theoretically loaded problem assumes that *de re* thought is object--dependent: the possibility of having the thought depends upon the existence of the particular object it is a thought about. Consider Rosie's singular thought about the book on her desk. Let w be a world that differs from actuality only in that one of the book-candidates does not exist in w. Surely the character of Rosie's thought in w is just as it actually is; for all the other candidates exist, and nothing in her relations to the books distinguishes that particular candidate from all others. Rosie's singular thought is not dependent on that candidate. Generalising: Rosie's singular thought is not dependent on any candidate. So Rosie's singular thought about the book is not object-dependent. At the very least, an alternative theoretical characterisation of singular thought is required.

1.5 Conclusion

This chapter presented both Unger's and Lewis's versions of the Problem of the Many. We began with two versions of Unger's argument. The first is a positive argument from a principle of minute differences to many cats on Tibbles's mat. We saw how to formulate this with very weak ontological assumptions, and that the puzzle is not primarily about the existence of individuals, but the instantiation of ordinary sortal properties. The second version is best seen as a challenge to our ordinary belief in just one cat on Tibbles's mat: what ensures that each macroscopic object is correlated with a unique class of microscopic constituents? Given this second version, rejection of the principles of minute differences doesn't solve the problem. We then turned to Lewis's puzzle. This proceeds by appeal to vagueness and borderline cases of parthood. Again, no controversial assumptions seemed to be required. Once Unger's puzzle of too many candidates and Lewis's puzzle of borderline candidates were in place, we saw that the question of whether these are two puzzles or one cannot be settled in isolation from a theory of vagueness.

Both these puzzles seem to arise for all sorts of ordinary macroscopic object, including ourselves. So we closed with a range of more and less theoretical reasons to be dissatisfied with simply accepting the conclusion that there are many people sitting in each of our chairs. The next chapter develops a supervaluationist account of vagueness. The final two chapters apply this account to the Problems of the Many in two different ways.

Chapter 2

Supervaluationist Theories of Vagueness

This chapter develops a supervaluationist approach to vagueness. The following two chapters present different applications of this approach to Unger's and Lewis's puzzles.

Although prominent in the literature on vagueness, supervaluationism is not a unified theory of vagueness. It is, rather, a collection of views united by a formal framework and the importance of the concept of super-truth to the analysis of clarity. The three key theses are:

(i) Vague classification is best represented by a class of sharpenings.

(ii) The apparatus of truth-evaluation privileges no sharpening over any other.

(iii) Clear truth is best represented by supertruth, and clear falsity by superfalsity.

§2.1 describes the formal setting and key concepts. §2.2 begins to provide the formalism with a philosophical interpretation. A range of interpretations are assessed in §§2.3–2.4, where all bar two are rejected. These remainders are the focus of the rest of the chapter. §2.5 presents supervaluationism's most attractive features. We turn to a range of objections in §§2.6–2.9. One view will emerge as clearly preferable to the other. Subsequent chapters apply this view to Unger's and Lewis's puzzles.

Two questions before we begin: (a) why supervaluationism?; (b) why *only* supervaluationism? In response to (a): because supervaluationism is popular, and maybe even the standard approach to vagueness insofar as there is such a thing. In response to (b): because there simply isn't space here to examine more than one approach to vagueness in the detail it deserves.

2.1 Supervaluationist formal theory

This section outlines the supervaluationist formalism. The classic presentation is Kit Fine's (1975). Because one of Fine's primary goals is to survey the formal terrain, his discussion contains more complexity than we require. So we simplify where possible. In particular, we consider only complete sharpenings that assign a truth-value to each wff, and not also partial sharpenings that relax this constraint.

Our object-language has the form of standard predicate calculus with identity. We use the following metalinguistic variables (alongside subscripted, primed and starred variants): 'α' ranges over object-language terms and variables; 'Φ' ranges over object-language predicates (sometimes superscripted to indicate the number of argument positions); 'A', 'B' and 'C' range over object-language wffs; 's', 't' range over sharpenings; 'v' ranges over variable assignments. We won't always mark use/mention distinctions, and will sometimes use metalinguistic variables schematically. Context should make things clear enough.

A *sharpening* s is a pair $\langle D_s, [\![]\!]_s \rangle$. D_s is a set of individuals, the *domain* of s.[1] $[\![]\!]_s$ is a valuation function from object-language constant terms and predicates such that:

For each term $\alpha : [\![\alpha]\!]_s \in D_s$.

For each n-place predicate $\Phi^n : [\![\Phi^n]\!]_s \subseteq D_s^n$.[2]

$[\![=]\!]_s = \{\langle x, x \rangle : x \in D_s\}$.

[1] We call the domain of s the s-domain. Similarly, a predicate's extension at s is its s-extension, a sentence true at s is s-true, and so on.

[2] For simplicity, we often write as if the members of $[\![\Phi^1]\!]_s$ were elements of D_s, rather than their singletons.

By way of initial gloss, $[\![\Phi]\!]_s$ is the extension of Φ at s: the set of things of which Φ is s-true. Some interpretations of the formalism mandate revisions to this gloss. The last of these conditions ensures that identity is classical.

A *supervaluationist model* M is a class of sharpenings such that:

For any $s, t \in M$ and singular term $\alpha : [\![\alpha]\!]_s = [\![\alpha]\!]_t$.

For any $s, t \in M : D_s = D_t$.

The first condition ensures that singular terms are not a source of vagueness. This is relaxed in the next chapter (§3.1.1). The second condition ensures that quantification is not a source of vagueness.

A *variable assignment* v is a function from object-language variables α such that:

$v(\alpha) \in D_s$.

Let $[\![]\!]_{s,v}$ be the function from object-language terms and variables such that:

For each term $\alpha : [\![\alpha]\!]_{s,v} = [\![\alpha]\!]_s$.

For each variable $\alpha : [\![\alpha]\!]_{s,v} = v(\alpha)$.

$[\![\alpha]\!]_{s,v}$ is the value of α given (i) the assignment $[\![]\!]_s$ of values to constant terms, and (ii) the assignment v of values to variables.

We use this to recursively define a relation $\Vdash$ between assignments v, sharpenings s, models M and wffs A, written '$v, s, M \Vdash A$'. When $v, s, M \Vdash A$, we say that A is true at v, s, M, or that v, s, M makes A true. We say that v, s, M makes A false iff v, s, M makes $\neg A$ true. Given the clause for $\neg$ below: v, s, M makes A false iff v, s, M does not make A true. Thus we can speak of the truth-value of a wff relative to a model, sharpening and assignment (though reference to a model will often be left tacit). Although the relativisation to models is inert in the definitions below, it's needed to introduce a clarity operator Δ later. We treat $\vee, \rightarrow, \exists$ as defined from $\neg, \wedge, \forall$ in the standard way. The base clauses of the definition are:

$v, s, M \Vdash \Phi^n \alpha_1, \ldots, \alpha_n$ iff $\langle [\![\alpha_1]\!]_{s,v}, \ldots, [\![\alpha_n]\!]_{s,v} \rangle \in [\![\Phi^n]\!]_s$.

$v, s, M \Vdash \neg A$ iff $v, s, M \nVdash A$.

$v, s, M \Vdash A \wedge B$ iff $v, s, M \Vdash A$ and $v, s, M \Vdash B$.

$v, s, M \Vdash \forall x A$ iff $v', s, M \Vdash A$, for every assignment v' that differs from v at most over 'x'.

Formally, sharpenings are classical models.

We now drop the relativisation to assignments:

$s, M \Vdash A$ iff $v, s, M \Vdash A$, for all assignments v.

Then we define supertruth and superfalsity in a model:

A is *supertrue* in M iff, for any sharpening $s \in M : s, M \Vdash A$.

A is *superfalse* in M iff, for any sharpening $s \in M : s, M \nVdash A$.

A is supertrue (superfalse) in M iff every sharpening in M makes A true (false). Thus we can talk of the supertruth-value of a wff in a model.

Now the formalism is in place, let's apply it to vagueness.

2.2 Understanding the supervaluationist formalism

The intuitive inspiration for supervaluationism is the idea that vague predicates can be made precise in many different ways; hence the interest in classes of sharpenings. If a predicate F applies to an object o regardless of how F is made precise, then o is clearly F. If F never applies to o, regardless of how F is made precise, then o is clearly not F. And if F applies to o under only some ways of making F precise, then o is borderline F. Hence the identification of clear truth with supertruth and clear falsity with superfalsity.

This is only a sketch of a guiding picture. What exactly is the sense in which a vague predicate can be made precise? What is a sharpening? These questions are best addressed in the context of attempts to delimit a consequence relation underlying vague discourse. There are two reasons for this. Firstly, focus on consequence pins down the relevant notions of content and semantics, thereby helping to eliminate terminological disputes: our interest is in logically relevant content. We'll see that this constrains permissible accounts of sharpenings and supervaluationist models. Secondly, because of the Sorites, an account of good deductive inference within a vague language is arguably the most pressing demand on any theory of vagueness; and a language's consequence relation provides the standard of correctness for deductive inferences within it. Thus before we can begin with a philosophical account of the formalism, an account of the relationship between model-theory and consequence is required. This is our next topic. Once this account is in place, §§2.3–2.4 use it to develop two kinds of supervaluationism. These are evaluated in the rest of the chapter.

2.2.1 Consequence, truth and interpretations

Alfred Tarski (1936) offered the following analysis of consequence:

> C is a consequence of Γ iff, for any interpretation I, if every member of Γ is true under I, then C is true under I.[3]

Throughout, 'Γ' ranges over sets of sentences. An interpretation is an assignment of logically relevant content to linguistic items, a possible semantic structure. So on the Tarskian view, the members of Γ jointly imply C iff there's no way of assigning content to the members of Γ and to C that makes the former true without also making the latter true.[4]

John Etchemendy (1990) criticises Tarski. Here are two examples. (i) were there only one thing, '$\exists x \exists y x \neq y$' would be true on no interpretation and hence, by Tarski's lights, be logically false. (ii) In higher-order languages, either the Continuum Hypothesis or its negation is true on all interpretations and hence logically

[3] It's debatable whether this is quite what Tarski intended. It's certainly how Etchemendy (1990) interprets him. But Tarski's own presentation is in terms of *models*, i.e. mathematical structures, rather than interpretations. Plausibly however, Tarksi's models were intended as mathematical representations of interpretations.

[4] Restrictions on which expressions get re-interpreted are required; for $A \wedge B$ wouldn't imply A if interpretations of $\wedge$ as disjunction were permitted. Our presentation builds the interpretations of $\wedge$, $\neg$, $\vee$, $\rightarrow$ $\forall$ and $\exists$ into the rules governing truth-evaluation. The target is a notion of formal consequence, where form is determined by which expression's interpretations are held fixed. We also need to insist on uniform interpretations, otherwise A wouldn't imply A.

true, despite mathematics plausibly not being part of logic.[5] Following Stewart Shapiro (1998), we therefore insert 'necessarily' between the 'iff' and quantifier over interpretations.[6] This avoids both kinds of problem, whilst remaining broadly Tarskian in spirit. On this approach, Γ implies C iff the members of Γ and C cannot be interpreted to make the former true and the latter false.[7]

The goal of model-theory is a mathematical representation of the space of possible interpretations. We want to define a class of mathematical structures S and a relation R between members of S and sentences A such that, for any possible distribution π of truth-values across sentences, there is a structure $x \in S$ such that: $R(x, A)$ iff $\pi(A) = \text{True}$. The structures represent interpretations. The condition under which $R(x, A)$ represents the truth-condition that A would have were the interpretation(s) represented by x the actual interpretation.

Model-theory cannot be completely successful. Etchemendy's example of the Continuum Hypothesis provides one example why. Unrestricted quantification provides another. In each model, the quantifiers range over a set. Since there is no set of all sets, model-theory cannot capture quantification over all sets. Still, these limitations shouldn't matter for our purposes. Our primary interest is vagueness, not unrestricted quantification or the outer reaches of logical possibility.

In order to use the supervaluationist formalism for this purpose, two things are required. First, we need to identify a class of elements of the formalism with the class of possible interpretations.[8] Second, we need to identify a dyadic relation R with the *true-under* relation between sentences and possible interpretations I: A is true under I iff A would be true if it had the content assigned it by I. We can then identify the truth-condition of A under I with the defined condition under which R holds between A and the representative of I.

[5] The reason for this is that the Continuum Hypothesis may be formulated in a second-order language that lacks non-logical vocabulary.

[6] In the relevant sense of necessity, it's possible for a sentence to be interpreted thus-and-so iff the language's semantic/compositional structure doesn't rule out it's being interpreted thus-and-so.

[7] The key thesis of model-theoretic semantics is that there are enough mathematical structures to represent every possible way of interpreting a language. Etchemendy's criticisms show that this assumption is false, and hence that there are representational limits on mathematised semantics.

[8] This is the sense of 'identify' in which identification is a form of representation: the items identified with interpretations are used to represent interpretations.

Regarding the first task, there are two candidate accounts of interpretations.

(i) Supervaluationist models.

(ii) Sharpenings.

§2.3 and §2.4 consider these in turn, and the result of combining them with various candidate accounts of *true-under*. All bar two combinations of views will be rejected. These are assessed in the remainder.

2.3 Interpretations as supervaluationist models

This section examines the identification of interpretations with supervaluationist models. On this view, each vague language has a unique semantic structure, represented by a supervaluationist model. Vagueness is a feature of an expression's content; a feature of the propositions expressed by sentences featuring that expression. Thus clear truth (falsity) becomes a semantic feature of propositions, due to the supervaluationist account of it as supertruth (superfalsity):

> M makes A clearly true iff A is supertrue in M.
>
> M makes A clearly false iff A is superfalse in M.
>
> M makes A borderline iff A is neither supertrue nor superfalse in M.

We want to add an account of consequence to this. So we need to convert the triadic $s, M \Vdash A$ relation into a dyadic $M \Vdash A$ to represent the *true-under* relation between interpretations/supervaluationist models and wffs. There seem to be three options:[9]

Particular $M \Vdash A$ iff $a, M \Vdash A$, where a is a privileged sharpening in M.

Subtruth $M \Vdash A$ iff, for some sharpening $s \in M : s, M \Vdash A$.

Supertruth $M \Vdash A$ iff A is supertrue in M.

Only the last of these stands up to scrutiny. We take them in turn.

[9] Although there are other possibilities, this isn't the place for an exhaustive survey. These are the most obvious and popular candidates.

2.3.1 Particular

The Particular account identifies truth with truth at a privileged sharpening in the intended interpretation. This conflicts with our initial characterisation of supervaluationism (p.64) by privileging one sharpening over all others when evaluating for truth. It would be less misleading to identify interpretations with the privileged sharpenings themselves, rather than supervaluationist models. (The result is formally akin to Williamson's (1994) epistemic view, described in §2.4.1.) So let us set this option aside.

2.3.2 Subtruth

On this view, a sentence is true iff true under some way of making its constituents precise. Dominic Hyde (1997) endorses this.

Consider a ball b which is a perfectly balanced red/orange borderline case. Both the following are borderline, and hence neither supertrue nor superfalse:

$$b \text{ is red}$$
$$b \text{ is orange}$$

Since neither is superfalse, each is true at some sharpening. So both are true *simpliciter*. But red and orange are incompatible. More generally, whenever A is borderline/supertruth-valueless, both A and $\neg A$ will be true. Yet that's logically impossible if $\neg$ expresses negation. The background picture is one on which the semantic rules overdetermine truth-value in borderline cases, and the result is inconsistency.

A response is available. Both 'b is red' and 'b is orange' can be supertruth-valueless in a model without both being true at the same sharpening. So both can be borderline without their conjunction 'b is red $\wedge$ b is orange' being true at some sharpening, and hence without being true. So it does not follow that their conjunction will be true. The defender of Subtruth may therefore claim to respect the incompatibility of red and orange by permitting no sharpening that places anything in the extension of both 'red' and 'orange'. (These penumbral connections are discussed in §2.5.4.)

Similarly, no sharpening makes both A and $\neg A$ true. So even if these sentences are both supertruth-valueless (and hence true), their conjunction $A \wedge \neg A$ will be superfalse and hence false. The defender of Subtruth may therefore claim to avoid inconsistency by making no contradictions true.

These responses are unsatisfactory. They amount to observing that if truth is subtruth, then Conjunction Introduction is unsound. One problem is that that principle is too central to our understanding of conjunction to give up. Another is that it doesn't address the initial problems: b has incompatible colours; both A and $\neg A$ are true. The response blocks expression of these problems using conjunction. But we can still truthfully say "b is red. b is orange." and $\ulcorner A.\ \neg A \urcorner$. These are no less objectionable than the conjunctions in question; endorsing incompatible claims using successive successive sentences is no better than doing so using a single conjunctive sentence. We should therefore reject the Subtruth account of truth.

2.3.3 Supertruth

This leaves the identification of truth with supertruth. Since a sentence is false iff its negation is true, we also have the identification of falsity with superfalsity. This is probably the most popular form of supervaluationism. Fine (1975), Hartry Field (1974), Vann McGee and Brian McLaughlin (1994; 2000), and Rosanna Keefe (2000) all endorse it in one form or another.

Since a sentence can be neither supertrue nor superfalse in M, by being true at only some sharpenings in M, this view violates the classical semantic principle of:

Bivalence For any sentence A, either A is true, or A is false.

Since clear truth and truth are both identified with supertruth, it follows that all borderline sentences lack truth-value. Each borderline sentence is a counterexample to Bivalence.

Plugging this view about truth into the Tarskian analysis of consequence gives the relation Williamson (1994, p.148) calls *global consequence*:

> $\Gamma \models_{\text{global}} C$ iff, for any model M, if every member of Γ is supertrue in M, then C is supertrue in M.

Consequence is preservation of truth-at-all-sharpenings in all models. Since sharpenings are classical models, all and only the classical logical truths are true at all sharpenings. Hence all and only the classical logical truths are $\models_{\text{global}}$-logical truths. So even though Bivalence is false, the following classical logical law is sound:

Law of Excluded Middle Every sentence $\ulcorner A \vee \neg A \urcorner$ is a theorem.

In fact, classical consequence and $\models_{\text{global}}$ coincide within predicate calculus. (Fine, 1975, p.125 sketches the proof.) Matters are more complex in languages enriched with a clarity operator Δ, but we'll come to that later (§2.6).

On the present view, each supervaluationist model represents an interpretation of a vague language. Can we say anything more? What, for example, is the interpretation of a predicate? Focus on monadic predicates for simplicity. What is a monadic predicate's contribution to truth-conditions, according to the present view?

In classical semantics, each interpretation I assign a set of objects to each predicate Φ, its I-extension. Φ's I-anti-extension is the set of objects (in the domain) that don't belong to Φ. Atomic predications are interpreted using the following truth- and falsity-conditions:

> $\Phi\alpha$ is I-true iff the I-referent of α belongs to the I-extension of Φ.
>
> $\Phi\alpha$ is I-false iff the I-referent of α belongs to the I-anti-extension of Φ.

Now, supervaluationist models are unlike classical interpretations because they don't actually assign any extra-linguistic semantic values to expressions: content is determined by all the sharpenings in a model, rather than assigned to expressions directly by supervaluationist models themselves. Nonetheless, we might seek to generalise the classical picture as follows.

Let $s, \ldots, t$ be every sharpening in M. Define the M-extension of Φ as the intersection of $[\![\Phi]\!]_s, \ldots, [\![\Phi]\!]_t$: the set of objects every sharpening assigns to Φ. Define the M-anti-extension of Φ as the set of objects $x \in D_s$ such that $x \notin [\![\Phi]\!]_s, \ldots, x \notin [\![\Phi]\!]_t$: the set of objects no sharpening assigns to Φ. Now we can offer the following account of atomic predications:

$\Phi\alpha$ is M-true iff the M-referent of α belongs to the M-extension of Φ.[10]

$\Phi\alpha$ is M-false iff the M-referent of α belongs to the M-anti-extension of Φ.

Given the definitions of M-extension and M-anti-extension, $\Phi\alpha$ is M-true (M-false) iff $\Phi\alpha$ is supertrue (superfalse) in M. This account therefore respects the identification of truth (falsity) with supertruth (superfalsity). The view departs from the classical picture because a predicate's anti-extension is not determined logically, solely on the basis of its extension. However, the classical conceptions of predication and of a predicate's contribution to truth-conditions are retained.

This view is problematic. Sharpenings play no role in its truth-conditions for atomic predications, yet they are indispensable to those for molecular statements. The semantics of atomic and molecular statements is therefore non-uniform. Since a uniform semantics is preferable, we should reject this account of the semantic role of predicates and of the truth-conditions of atomic predications. Instead, a predicate's semantic contribution should be identified with its role in delimiting the space of sharpenings as a whole. On this view, semantic relations between expressions prevent the attribution to them of discrete semantic contributions.

The question now arises: what is a sharpening? The following five sections address the following five answers in turn. Sharpenings are...

...ways a vague language could be made precise.

...ways a precise boundary could be drawn.

...classical interpretations.

...theoretical posits.

...artefacts of the formalism.

We will eventually settle on the last of these.

[10] M-reference is unproblematic because we imposed the following constraint on models M: $[\![\alpha]\!]_s = [\![\alpha]\!]_t$, for all sharpenings $s, t \in M$ and singular terms α.

2.3.3.1 Sharpenings as ways a vague language could be made precise

This is perhaps the most natural account of the supervaluationist formalism. It seems to be endorsed by Keefe (2000, p.154), a prominent supervaluationist.[11] The idea is that a vague language can be made precise in many different ways. If every way of making the language precise makes A true, then A is clearly true; if every way makes A false, then A is clearly false; and if some ways make A true while others make A false, then A is borderline.

This account of sharpenings faces two problems. The first is that it confuses counterfactual semantic status with actual semantic status. This will not move advocates of the view, however. For their core thesis is that actual truth *is* truth in every possibility where the language is made precise.

The second problem arises from the fact that, "[t]o make an expression precise, uncontroversial truths involving it must be preserved" (Keefe, 2000, p.154). This is the key difference between making a meaning precise and replacing it with a precise meaning. But as Jerry Fodor and Ernest LePore (1996) observe, sharpenings do not respect all uncontroversial truths. Consider:

Everyone greater than 5'11" in height is tall

Each way of making 'tall' completely precise makes this true or makes it false. But since it's analytically borderline, and hence neither true nor false, no way of making 'tall' completely precise respects the meaning of 'tall'. Consider also:

It's borderline whether everyone greater than 5'11" in height is tall

Each way of making 'tall' completely precise makes this analytic truth false. So no way of making 'tall' completely precise respects every uncontroversial (analytic) truth involving it. So sharpenings are not ways a vague language *could* be made completely precise.

Keefe replies that...:

> "...this objection misrepresents the role of precisifications: such valuations do indeed fail to capture all features of the meanings of our pred-

[11] Keefe doesn't carefully distinguish accounts of sharpenings. We'll see in §2.3.3.3 that she endorses different accounts on different pages.

> icates. . . . But this constitutes no *objection* to the theory, for the claim is that it is quantification over all precisifications that captures the meaning of the natural language predicates; the individual precisifications need not.” (Keefe, 2000, p.190; original emphasis)

This response fails. The objection wasn’t to the identification of truth with supertruth *per se*, but to combining that identification with the present account—indeed, Keefe’s own account—of sharpenings. Simply reaffirming this combination of views does not make them consistent. So we should reject this account of sharpenings.

2.3.3.2 Sharpenings as ways a precise boundary could be drawn

This option explains sharpenings in terms of boundary-drawing. Sometimes we must decide whether to count a borderline case as a positive or negative case. Suppose the owner of a record shop is re-organising her stock by genre. She encounters a tricky case: should the John Adams records go in the minimalism section? A decision is required one way or another, but competence with ‘minimalism’ enforces neither choice. Suppose the store owner decides not to count Adams amongst the minimalists. This has consequences for her classification of the remaining stock: nothing less minimal than Adams counts as minimalism.[12] There seems nothing illegitimate about this commonplace aspect of linguistic practice.

The proposal is to treat sharpenings as formal representations of the effect of such decisions on classification. Borderline cases are those that can be competently called either way; that’s all there is to being a borderline case.

Three problems arise. Firstly, it is doubtful whether we can make classificatory decisions that settle all borderline cases. Secondly, competence arguably mandates leaving a classificatory gap, however small (Shapiro, 2006, pp.8–12). Thirdly, §2.7.2 argues that if borderline sentences lack truth-value, then borderline classificatory decisions are semantically illegitimate. If so, then sharpenings cannot be explained in of legitimate such decisions. So we should reject this account of sharpenings.

[12] Similar issues arise for ‘less minimal than’.

2.3.3.3 Sharpenings as classical interpretations

This view holds that a range of classical interpretations all fit the meaning-determining facts equally well. None is privileged over any other as the actual, or *intended*, interpretation of our language. Rather, these interpretations all contribute jointly to the determination of (super)truth-conditions. Field (1974) and Keefe (2000, pp.155–9) endorse views along these lines. Keefe also claims that it fits the picture of vagueness as "semantic indecision" endorsed by Lewis (1986b, p.212; 1993a).[13] We will see that Keefe is wrong about this.

This view can explain the supervaluationist formalism in terms the classical semanticist finds legitimate. It departs from classicality by first generalising the classical metasemantic[14] picture to accommodate meaning-determining facts too coarse-grained to rule out all bar one classical interpretation. Semantic departure from classicality then comes from identifying truth with truth on all intended classical interpretations.

The coherence of this view is doubtful. It involves two theses: (i) the meaning-determining facts don't determine a unique intended interpretation; and (ii) each vague language possesses a unique intended interpretation. Taken at face value, these are obviously inconsistent. The inconsistency is supposed to be resolved by taking the interpretations in (i) as classical interpretations and those in (ii) as supervaluationist models. Two problems arise. The first is that the metasemantic concepts used to determine the classical interpretations enter into the truth-conditional content of all ordinary vague expressions, if this view is correct. The second is that determination of a unique intended vague interpretation (supervaluationist model) should suffice to ensure that no precise (classical) interpretation fits our use of language even approximately: if our use of language determines a vague content, then no non-vague content is even a candidate content. The lesson of this second point is that if truth is supertruth, then vagueness is not semantic

[13] Note that Lewis does not combine the semantic indecision view the claim that truth is supertruth. His (1970a, p.228) does propose that identification, but explicitly not in combination with the semantic indecision picture.

[14] Metasemantics is the study of how expressions come to possess semantic properties, and what those properties are.

indecision.

On this account of sharpenings, supertruth is more naturally seen as a partly semantic and partly metasemantic concept, rather than as the primary notion of semantic evaluation, i.e. truth. §2.4 develops this kind of view. We set aside the account of sharpenings as classical interpretations until then.

2.3.3.4 Sharpenings as theoretical posits

This view suggests that sharpenings should be treated like any other theoretical posit, and 'sharpening' like any other theoretical term. We can make this precise using a variant on Lewis's (1970b) account of theoretical terms.

Let S be the supervaluationist theory (as formalised in predicate calculus). Let $S(X)$ be the result of replacing every occurrence of 'is a sharpening' in S with the unbound predicate-variable 'X' (where 'X' does not occur in S). Then the property of *being a sharpening* is whichever (unique) property satisfies $S(X)$. The supervaluationist theory S thereby provides an implicit definition of what it is to be a sharpening. Sharpenings are like electrons in this respect: although 'electron' isn't explicitly definable using everyday vocabulary, electron-theory permits an implicit definition of the electrons as whichever objects behave as it claims.

This leaves us with only descriptive knowledge of sharpenings, not particular knowledge: the sharpenings are whatever occupy the sharpening-role in supervaluation-theory. Our knowledge of sharpenings is just like your knowledge of quarpenings, if all you know about quarpenings is: quarpenings are what Nick has in his pocket. This is no objection to the view. For sharpenings are no different in this respect than any other theoretical posit. Furthermore, there are reasons to think that a significant portion of our ordinary knowledge is also descriptive (Lewis, 2009).

On this view, belief in sharpenings is justified to the extent that supervaluationism as a whole is adequate. The better it can accommodate our use of vague language, the better the justification for believing in sharpenings. The key difficulty for the view is that we can identify truth with supertruth without incurring ontological commitment to sharpenings or compromising supervaluationism's explanatory ambitions. We examine this rival next.

2.3.3.5 Sharpenings as artefacts

This final view denies that sharpenings exist, other than as the pure mathematical subjects of the supervaluationist formal theory. McGee and McLaughlin (1994, 2000) and Josh Dever (2009) endorse views along these lines. Supervaluationism is taken as a formal framework structurally similar to some aspects of vague language, though not necessarily to all. Although some features of the framework correspond to features of vague language, they need not all do so; and the present view claims that sharpenings don't.

Roy T. Cook (2002) and Shapiro (2006, ch.2), distinguish three attitudes towards applied mathematical theories:

Representationalism Every feature of the formalism represents a feature of the target system.

Modelling Some, but not necessarily all, features of the formalism represent features of the target system.

Instrumentalism No features of the formalism represent features of the target system (only input-output matching and predictive success matter).

Representationalism and Instrumentalism lie on extreme ends of a spectrum of views. Both extremes are problematic.

Representationalism makes very high demands on the deployment of formal tools. We want a mathematical representation of all possible assignments of logically relevant content to a vague language. But there is no pre-theoretic reason to expect that any mathematical structure will exactly resemble the structure of a vague interpretation, especially since the structures in question have to be comprehensible to, and manipulable by, beings like ourselves. §2.9.6 even argues that no mathematical structure will do so. So if Representationalism is correct, there is no reason to expect that our project will be a success. Even the slightest departure from perfection would render it a total failure.

Instrumentalism is problematic because we would like more than mere empirical adequacy and predictive success from our theorising. We would like to know why a successful theory is successful. The most straightforward explanation is that (at least some of) the theory's internal structure corresponds to structure in the target system. But that explanation is incompatible with Instrumentalism.

Modelling provides a moderate alternative to Representationalism and Instrumentalism, that avoids their worst excesses whilst accommodating their key insights. Those aspects of the formalism that represent features of the target system we call *representors*; those that don't we call *artefacts*. Can we say anything general and informative about which features of which theories fall into which category? Well, our objection to Instrumentalism suggests that features essential to a theory's explanatory and predictive success shouldn't be treated as artefacts. And the following version of Ockham's Razor suggests that, *ceteris paribus*, only those features should be treated as representors:

> Posit as few kinds of entities as are necessary to explain a theory's success.

It follows that exactly those features of the formalism necessary to explain a theory's success should be treated as representors. Are sharpenings amongst those features of supervaluationism? Arguably not.

Sharpenings are needed to formulate tractable and comprehensible definitions of the relations between truth-conditions which interpret the connectives. But they need serve no other role. We can regard them as mere calculating devices used to determine the truth-conditions of wffs, rather than as components of the semantic structures of vague languages. McGee and McLaughlin (1994, §4) and Dever (2009, §6) even offer accounts of the theoretical utility of sharpenings in terms of features of the consequence relation for a vague language.

From the point of view of representational content, the resulting view treats a supervaluationist model as a black box: information about the world goes in (in the form of information about membership), and truth-values come out. Vague interpretations either lack internal structure, or the supervaluationist formalism does not capture it. Given the difficulties with the other accounts of sharpenings, we will henceforth assume that if interpretations are supervaluationist models and truth is supertruth, then sharpenings are artefacts of the formalism.

2.3.4 Interpretations as models: concluding remarks

This section examined the identification of interpretations with supervaluationist models. Views that don't identify truth with supertruth were rejected in §§2.3.1–2.3.2, and all bar one account of sharpenings were rejected in §§2.3.3.1–2.3.3.5. This leaves only the view that combines these three theses:

(i) Interpretations are supervaluationist models.

(ii) Truth under an interpretation is supertruth in a supervaluationist model.

(iii) Individual sharpenings don't represent anything in the semantic structures underlying vague language.

The conjunction of (i)–(iii) we call the *Supertruth View*. Of all the views that endorse (i), it is the only one we will consider in the remainder (although the following discussion won't turn on whether sharpenings are treated as artefacts or theoretical posits). We now examine a different account of interpretations.

2.4 Interpretations as sharpenings

This section develops a view that identifies interpretations with sharpenings, rather than with supervaluationist models. Vagueness is located in a language's association with a range of classical semantic structures (represented by the members of a supervaluationist model). §§2.4.1–2.4.2 develop the metaphysics. §2.4.3 turns to truth and consequence. §2.4.4 responds to an objection.

2.4.1 The association relation

The view is under-specified without an account of the association relation between languages and classes of interpretations. What is it for a language to be associated with many interpretations? What is the representational role of supervaluationist models? By way of illustration, this section outlines Williamson's (1994) epistemic view.

According to Williamson, each vague language possesses a unique intended classical interpretation: vague sentences express unique classical propositions. But there are many possibilities indiscriminable (to beings like ourselves) from actuality and in which the intended interpretation of our language is not its actual interpretation. For all we know, these possibilities could be actual; each of these counterfactually intended interpretations could actually be intended. The borderline cases are those over which these interpretations disagree. Vagueness thus becomes a form of semantic ignorance.

Williamson's epistemicism offers a philosophical account of the supervaluationist formalism. Supervaluationist models represent epistemic states of typical language users. The sharpenings within a model represent the interpretations that typical speakers cannot know to be incorrect interpretations of their language; a typical speaker's true belief that one or other is the actual intended interpretation could easily have been wrong. Clarity, as represented by supertruth, is thereby analysed using both semantic and epistemic concepts: x is clearly F iff our imperfect semantic knowledge doesn't prevent us from knowing that x is F. The relativisation of $\Vdash$ to models as well as sharpenings allows expression of claims whose truth-value depends on the epistemic states of typical speakers as well as on the interpretations of their expressions (though the expressive resources to do so must await the clarity operator Δ introduced in §2.5.1).

This account of the formalism conflicts with our initial characterisation of supervaluationism by privileging one sharpening over all others in the determination of truth-value. It does however, highlight the need for more detail about the representational role of supervaluationist models before we have a fully-fledged theory of vagueness. The following account of this missing detail draws inspiration from Lewis's brief and scattered writings on vagueness.

2.4.2 A Lewisian theory of association

In "General semantics", Lewis writes:

> "[W]e have so far been ignoring the vagueness of natural language. Perhaps we are right to ignore it, or rather to deport it from semantics to the

> theory of language-use. We could say, as I do elsewhere [Lewis (1969, ch.5)], that *languages themselves are free of vagueness* but that the linguistic conventions of a population, or the linguistic habits of a person, select not a point but a fuzzy region in the space of precise languages."[15] (p.228 Lewis, 1970a, my emphasis)

Lewis's languages are our interpretations.[16] Elsewhere he adds:

> "Super-truth, with respect to a language interpreted in an imperfectly decisive way, replaces truth *simpliciter* as the goal of a cooperative speaker attempting to impart information. We can put it another way: Whatever it is that we do to determine the 'intended' interpretation of our language determines not one interpretation but a range of interpretations. (The range depends on context, and is itself somewhat indeterminate.) What we try for, in imparting information, is truth under all the intended interpretations." (Lewis, 1993a, p.172)

The idea is to analyse vagueness via multiplicity of intended interpretation.

The meaning-determining facts settle which interpretations bear on linguistic communication within a community of language-users. These are the intended interpretations of that community's language; they assign to its expressions the content that members of the community express when using those expressions to communicate with one another. We say that any such interpretation *fits* the community's use of the language, or simply fits the community for short.

Think of the meaning-determining facts as determining a triadic relation: *interpretation x fits community y at least as well as interpretation z*. This relation induces an ordering on interpretations relative to a given community. An intended inter-

[15] Lewis continues: "However, it might prove better to treat vagueness within semantics, and we could do so as follows." He then outlines a degree-theoretic version of the Supertruth View.

[16] Matters are a little more complex than this. Lewis's languages are functions from sentences onto truth-conditions. These languages are defined by grammars. A grammar G is a pair $\langle S_G, [\![]\!]_G \rangle$, where S_G is a function from sentences onto syntactic structures and $[\![]\!]_G$ is a function from the basic constituents of these structures onto semantic values. Unlike Lewis, we assume a fixed syntactic structure. Our interpretations are therefore the second elements of the class of grammars with the appropriate shared first element.

pretation is a greatest element in this ordering:[17] no other interpretation fits the community at least as well as it does. When the ordering is total, the community's language has a unique intended interpretation and their utterances express unique propositional contents. When the ordering is only partial however, there may be many greatest elements, each of which is an intended interpretation of the community's language.[18] Utterances by members of such communities therefore express multiple propositional contents.

This Lewisian approach to vagueness claims that vagueness is multiplicity of intended interpretation. Our use of vague language is too coarse-grained to determine a total ordering on interpretations. The result is that utterances made using our vague language expresses many propositional contents, so similar to one another that ordinary usage doesn't distinguish between them. A vague language has not one but many semantic structures, each of which fits the meaning-determining facts well enough to count as really giving the language's content, and none of which fits those facts better than any other.

We can now provide an account of the supervaluationist formalism. Models represent states of the meaning-determining facts. Sharpenings represent interpretations. A model M represents a sharpening s as a greatest element in the fit-ordering (i.e. as an intended interpretation) iff $s \in M$. Different models represent different ways the meaning-determining facts could bear on intended interpretations.

On this view, the semantics is classical. The departure from classicality is not semantic but metasemantic: vague languages are multiply interpreted, they have many intended interpretations instead of just one.

This provides an analysis of clarity. A sentence is clearly true (as used by a given community) iff true under every intended interpretation (of the community's language); clearly false iff false under every intended interpretation; and borderline iff

[17] Greatest elements must presumably also exceed some threshold in the fit-ordering in order to count as intended: intended interpretations are those that fit well enough, as well as better than any other.

[18] Think of a partial ordering as a branching tree-like structure. The intended interpretations are the terminal nodes in the tree corresponding to the partial fit-ordering.

true under some but not all intended interpretations. Clarity is thereby explained using both semantic and metasemantic concepts.

2.4.3 Truth and consequence

We've got an account of the representational role of supervaluationist models, sharpenings and the relation between them. We want an account of consequence. So we need an account of *true-under*. Models represent states of the metasemantic facts. So we want to convert the triadic relation $s, M \Vdash A$ into a dyadic relation $s \Vdash A$. The extension of this relation will represent the extension of *true-under*, as it would be were the metasemantic facts as M represents them. The options are analogues of views considered earlier:

Particular Suppose the metasemantic facts are as M represents them. Then: $s \Vdash A$ iff $s, M \Vdash A$.

Subtruth Suppose the metasemantic facts are as M represents them. Then: $s \Vdash A$ iff, for some sharpening $t \in M$, $t, M \Vdash A$.

Supertruth Suppose the metasemantic facts are as M represents them. Then: $s \Vdash A$ iff A is supertrue in M.

Plugging these into the Tarskian account of consequence result in different consequence relations.

Subtruth succumbs to the logical objections in §2.3.2. Supertruth succumbs to the objection in §2.3.3: it gives a non-uniform semantics for atomic and molecular statements. Furthermore, any account of truth under an interpretation s that involves reference to, or quantification over, interpretations other than s succumbs to another objection: the metasemantic concepts that delimit the space of sharpenings enter into the truth-conditional content of every vague sentence.

This leaves only the Particular account. Plugging this into the Tarskian analysis of consequence gives what Williamson (1994, p.148) calls *local consequence*:

> $\Gamma \models_{\text{local}} C$ iff, for any model M and sharpening $s \in M$, if $s, M \Vdash \Gamma$, then $s, M \Vdash C$.

'$s, M \Vdash \Gamma$' means that $s, M \Vdash A$, for all $A \in \Gamma$. $\models_{\text{local}}$ preserves truth under an interpretation, whatever the state of the metasemantic facts. This account of $\models_{\text{local}}$ is exactly analogous to the account of classical consequence in standard possible-worlds semantics for modal logic. $\models_{\text{local}}$ is therefore classical. Classical logic and semantics are preserved wholesale by this Lewisian approach.

2.4.4 An objection: monadic truth

This section considers an objection to this Lewisian account of vagueness. The view associates each vague language with a class of intended classical interpretations. A sentence is borderline iff true under some but not all of them. But since a sentence is true iff true under its intended interpretation, and false iff false under its intended interpretation, it follows that borderline sentences are both true and false, which is impossible. So the relativised 'A is true under s' cannot be converted into an un-relativised 'A is true *simpliciter*'. The objection is that the Lewisian cannot accommodate the monadic nature of truth.

As it stands, the objection is under-specified. It might concern either the monadic English predicate 'is true', or the monadic property of *being true* . Either way, a response is available. We take these disambiguations in turn.

2.4.4.1 The monadicity of 'is true'

A response to the first disambiguation will provide an account of the extension of the English 'is true'. The most natural suggestion is:

> Suppose the metasemantic facts are as M represents them. Then $s, M \Vdash \ulcorner A$ is true$\urcorner$ iff $s, M \Vdash A$.

On this account 'is true' s-applies to exactly the sentences true under s.[19] That predicate therefore expresses the primary property of semantic evaluation of English sentences. Since those sentences express many contents, there are many such properties (one for each correct of content expressed). It doesn't follow that there are many ways the world itself is because it's entirely language- and sharpening-

[19] We ignore Liar-like paradoxes for simplicity.

independent which truth-conditions are satisfied. It's just that since vague sentences have many contents, there are many equally correct ways of evaluating them.

2.4.4.2 The monadicity of *Truth*

This second disambiguation of the objection claims that our Lewisian proposal is incompatible with the monadicity of the property of *Truth*. We respond by distinguishing this property from the primary property by which sentences are evaluated. The Lewisian proposal claims that this latter property is a relational property defined in terms of the monadic property of *Truth*

Note first that sentences are just concrete objects, strings of sounds or marks. These concrete objects aren't intrinsically meaningful, and hence aren't intrinsically suitable for semantic evaluation either. Sentences become suitable objects of semantic evaluation only through their use by members of a linguistic community to communicate with one another. This use bestows content upon sentences.

The moral is that the primary truth-bearers are not sentences themselves, but the contents they express. The Lewisian account of vagueness makes no claims about contents and their evaluation, but only about sentences and their evaluation. It is therefore compatible with the monadicity of (propositional) *Truth*. Evaluating sentences is a three-step process. The first step identifies their propositional content(s). The second step evaluates those propositions for *Truth*. The third step assigns truth-values to sentences, depending on whether they express true contents. Sentential truth is thus a relational notion defined in terms of the non-relational notion of propositional truth and the expression of propositions by sentences. The Lewisian theory of vagueness simply claims that this expression relation may be many-many: many sentences may express the same proposition, and a single sentence may express many propositions. This prevents there from being a single uniquely correct way of assigning truth-values to sentences. But since the Lewisian view is compatible with their being a uniquely correct assignment of truth-values to propositions, this is unproblematic.

2.4.5 The story so far

Two accounts of vagueness are now in place. One is the Supertruth View from §2.3. This associates each vague language with a unique vague intended interpretation—viz. a supervaluationist model—at the cost of a non-classical semantics. The other is the Lewisian proposal from §2.4. This maintains classical semantics by associating vague languages with many intended interpretations. We will call this the *Sharpening View*. The rest of this chapter assesses their relative merits. The Sharpening View will emerge as clearly preferable. The next two chapters apply it to Unger's and Lewis's puzzles in two different ways.

2.5 Four benefits

Why should we believe either of our philosophical interpretations of the supervaluationist formalism? This section presents four key benefits the supervaluationist may claim: an analysis of clarity and borderline status (§2.5.1); a response to the Sorites (§2.5.2); compatibility with classical logic (§2.5.3); respect for penumbral connections between borderline cases (§2.5.4). Some of the problems for supervaluationism examined in subsequent sections (§§2.6–2.9) show that not all of these benefits are available on both the Supertruth View and Sharpening View; the Supertruth View will fare particularly badly in this respect.

2.5.1 Analysing clarity

Supervaluationism offers both formal and philosophical analyses of clarity. We present the formal analysis now, and the philosophical analyses below (§§2.5.1.1–2.5.1.2).

Supervaluationism identifies clear truth with supertruth, clear falsity with superfalsity, and borderline status with lack of supertruth-value. This allows introduction of a clarity operator Δ akin to the $\Box$ of standard possible-worlds semantics for modal logic:

$$v, s, M \Vdash \Delta A \text{ iff } v, t, M \Vdash A, \text{ for all sharpenings } t \in M.$$

ΔA formalises $\ulcorner$It is clearly the case that $A\urcorner$. This semantics for discourse about clarity and borderline status is both familiar and tractable. Of course, the value of this formal treatment is secondary to the philosophical insight it brings. But it is an attractive feature nonetheless. So, what philosophical account of clarity can the supervaluationist provide? That depends on whether the Supertruth or Sharpening View is in question.

2.5.1.1 Clarity and the Supertruth View

The Supertruth View identifies truth with supertruth and falsity with superfalsity. Given their account of clear truth as supertruth and clear falsity with superfalsity, it follows that truth is clear truth, falsity is clear falsity, and borderline sentences fall down a truth-value gap. Borderline Fs fail to be F and fail to be not-F. This amounts to an analysis of clarity using (relatively) familiar semantic concepts. §2.7.1, §2.7.2 and §2.9.10 raise doubts about this analysis. This section raises a different problem, followed by two attractive features of the account.

Borderline status is not the only potential source of truth-value gap. Other examples include reference-failure and incomplete stipulations. An adequate analysis of clarity must distinguish vagueness-related truth-value gaps from those arising from other phenomena, if there are any. The natural account is that only vagueness-related gaps result from cross-sharpening variation in truth-value. But since our preferred account of sharpenings treats them as artefacts, that explanation isn't available. Treating sharpenings as theoretical posits doesn't alleviate the problem, but only makes the explanation circular: cross-sharpening variation in truth-value is implicitly defined in terms of vagueness-related truth-value gaps. The best the Supertruth theorist can achieve, it seems, is to list all other sources of truth-value gap and claim: vagueness-related gaps result from none of these other sources of gap. This isn't entirely satisfactory. But let us simply bracket this worry for the time being and move on.

This account of vagueness explains two aspects of our use of vague language. The first is that, if A is borderline, then there's something improper about the unqualified assertion "A". The explanation is that in making an assertion (i) one incurs commitment to that which was asserted, and (ii) one ought to communicate information about the state of the world. When A isn't true, one who asserts "A" therefore (a) commits themselves to something that isn't true, and (b) violates a norm governing assertion by communicating misleading information. §2.7.2 argues that this explanation is too strong.

The second feature of vague language explained by this semantic analysis of clarity is the seeming misguidedness of investigating whether A when it's known that A is borderline. The explanation is that A's borderline status precludes there being a correct answer to the question of whether A by making neither A nor $\neg A$ true. Investigation into a known borderline claim A is misguided because that knowledge is incompatible with knowing whether A. §2.8 argues that this explanation fails.

2.5.1.2 Clarity and the Sharpenings View

The Sharpening View analyses clarity using a combination of semantic and metasemantic concepts: clear truth is truth under every intended interpretation; clear falsity is falsity under every intended interpretation; and borderline status is variation in truth-value across intended interpretations. Not all of a borderline sentence's propositional contents are true. This explains why we shouldn't make unqualified assertions using borderline sentences.

The previous section noted two related components to the practice of assertion: (i) commitment to what one asserts; (ii) the communication of information. Each component justifies the rule:

> Assert only the truth.

Sentences that aren't supertrue express false propositions. Those who make assertions using borderline sentences therefore (a) commit themselves to falsehoods,[20] and (b) communicate misleading information. Thus we have:

[20] Why must using A to make an assertion bring commitment to all of the propositions A expresses? Couldn't it simply be indeterminate which of those many propositions the commitment was to? The answer is that this blocks a non-circular analysis of vagueness via multiplicity of interpretation. Given this requirement, commitment to every proposition expressed by the sentence used to make the assertion follows from the idea that the meaning-determining facts don't distinguish one from amongst many (good enough) candidate intended interpretations.

Use A to make an assertion only if A is supertrue.[21]

Unqualified assertions made using borderline sentences are illegitimate because they violate this (derived) rule.[22,23]

This explains why "Yes" and "No" are equally inappropriate answers to the question "A?", when A is borderline. For answering "Yes" amounts to the assertion "A". And answering "No" amounts to the assertion "$\neg A$". Since neither A nor $\neg A$ is supertrue, neither is assertable. So neither "Yes" nor "No" is an appropriate answer to the initial question. We can use this to explain borderline ignorance and the misguidedness of investigation into known borderline claims.

There are two candidate necessary conditions on the s-truth of $\ulcorner S$ knows that $A\urcorner$:

The proposition s assigns to A is true.

Any proposition that any intended interpretation s' assigns to A is true.

The first ensures that when A is borderline, $\ulcorner S$ knows that $A\urcorner$ is no better than borderline. The second ensures that when A is borderline, $\ulcorner S$ knows that $A\urcorner$ is clearly false. Either way, $\ulcorner S$ knows that $A\urcorner$ is not supertrue, and hence not assertable when A is borderline. A's borderline status excludes knowledge whether A in the sense of making it in-principle illegitimate to claim to know that A and also illegitimate to know that $\neg A$. Known borderline status makes investigation misguided because that knowledge is incompatible with legitimately claiming to know the result of the investigation.

---

[21] Sentences can be true under varying proportions of intended interpretations. So one violation of this rule can be worse than another. Thus one borderline assertion can be in worse standing than another.

[22] Could the extension of 'asserted' vary across sharpenings in the following manner?: if S used A to make an assertion, then 'S asserted that p' is s-true iff A s-expresses that p. This makes it borderline whether, rather than clearly false that, one obeys the first rule above when one makes an assertion using a borderline sentence. But there remain intended interpretations on which one fails to adhere to the rule. So using borderline sentences to make assertions complies with the norms governing assertion less well than does using supertrue sentences to do so, though it complies better than does using clearly false sentences. This may explain why borderline assertions are better than clearly false ones.

[23] Dorr (2003, §1 and pp.104–5) presents a related argument for a similar conclusion.

2.5.2 The Sorites

The challenge of the Sorites is to explain why the seemingly valid argument from the following three apparent truths to contradiction is unsound, and also why it appears to be sound:

R1 Ten seconds after his birth, Russell was young.

R2 Forty years after his birth, Russell was not young.

R3 If Russell was young i seconds after his birth, then he was also young $i+1$ seconds after his birth.

R1 and R2 really are indubitable: unless 'young' is trivial, some premisses of these forms are true. The argument from R1–R3 to contradiction is classically valid and formalisable in predicate calculus, hence both $\models_{\text{global}}$- and $\models_{\text{local}}$-valid. The fault lies in R3: every complete sharpening makes some instantiation of R3 false.[24] Hence R3 is clearly false. So why does it seem compelling?

Note first that since it's vague when Russell ceased to be young, the location of the young/not-young distinction varies across sharpenings. Let '$\textsc{young}(x, y)$' formalise 'x was young y seconds after x's birth. Then no instantiation of the following is supertrue:

$$\exists i(\textsc{young}(\text{Russell}, i) \wedge \neg\textsc{young}(\text{Russell}, i+1)) \tag{1}$$

The Supertruth theorist may now offer the following conjecture:

> Typical speakers reason as if a true existential generalisation required a true instantiation.

This can't be quite right because not everything is named.[25] But we'll assume that everything is named because it simplifies exposition without affecting anything of substance. The Supertruth theorist's conjecture amounts to the claim that typical speakers reason classically with existentials. This has the attractive feature of treating them as reasoning in accord with a natural and simple (i.e. classical) semantics.

---

[24] We assume for simplicity that everything has a (clear) name. Nothing of substance turns on this.

[25] Our introduction to the Sorites in §1.2.1.1 contains discussion related to the inadequacies of this conjecture.

Assume that ordinary speakers detect that (1) lacks a true instantiation. The Supertruth theorist's conjecture then implies that typical speakers will regard (1) as false. Since the negation of (1) is equivalent to R3, the Supertruth theorist has an account of why R3 seems compelling: typical speakers infer it by an invalid (but natural) argument from the correct observation that no instantiation of (1) is true.

The Sharpening theorist can't quite adopt this account because they reject the identification of truth with supertruth: every true existential does have a true instantiation. They could offer an alternative conjecture:

> Typical speakers reason as if a true existential generalisation required a clearly true instantiation.

Since (1) lacks a clearly true instantiation, this will achieve the same result. But this alternative conjecture cannot be justified by attributing classical inferential practice to ordinary speakers. In effect, the conjecture treats speakers as mistaken about the content of their existentially quantified claims, but fails to explain why the mistake arises. Luckily for the Sharpening theorist, two alternative explanations are available.

The first explanation begins by observing that (1) lacks a clearly true instantiation. The previous section offered an account of why borderline status makes (unqualified) assertion illegitimate. So no instantiation of (1) is assertable. Given an argument from unassertability to falsity, the Sharpening theorist may then explain the apparent falsity of (1), and hence truth of R3, by attributing reasoning in accord with that argument to typical speakers. One such argument is as follows.

The unassertability of instantiations of (1) isn't the result of our own limitations. No amount of investigation into Russell's gradual aging could make any such instantiation assertable. But when unassertability isn't the result of our own limitations, then that's because there's no truth there to assert.[26] So every instantiation of (1) must be false. So (1) must be false. So R3 must be true. Attributing this line of argument to ordinary speakers treats them as ignorant (or forgetful) of

[26] A counterexample: no instantiation of 'some mammal was the first unnamed dog born at sea' is assertable (or even true), and not because of our own limitations; yet there's no temptation to regard it as false. However, since this contains semantic vocabulary within the scope of a quantifier, the Sharpening theorist may claim that it is relevantly disanalogous to (1).

another source of unassertability: the expression of multiple contents, only some of which are true.

The second explanation for the attraction of R3 begins by observing that Russell ages gradually. No single-second duration in his aging is more intrinsically significant than any other. Neither, the Sharpening theorist claims, do successive single-second durations stand in significantly different relations to our use of 'young'. R3 is a natural, though incorrect, way of reporting the following consequence of those two facts: no single-second duration marks any significance difference as to whether 'is young' applies to Russell, as that predicate is used within the community of speakers of English.

2.5.3 Classical logic

§2.3.3 and §2.4.3 noted that $\models_{\text{local}}$ and $\models_{\text{global}}$ coincide with classical consequence within predicate calculus. Hence both the Supertruth View and the Sharpening View preserve classical logic when reasoning within languages of that form. This doesn't however, extend to languages containing Δ. More on that shortly (§2.6).

Now, classical logic isn't absolutely mandatory. But it is (a) the default, (b) required for large portions of best science and mathematics, and (c) hard to see how it could fail. Any departure from classical logic had therefore better (i) have strong positive arguments in its favour, (ii) validate classical reasoning in mathematics, and (iii) provide insight into the underlying semantic features of the language responsible for non-classicality. Since supervaluationism doesn't satisfy (i) or (iii), it's a good thing that it preserves classical logic.[27]

2.5.4 Penumbral connection

Let *b* be a ball that's a perfectly balanced red/orange borderline case. It would be misleading to describe *b* without qualification as red, or to describe it without qualification as orange. Nonetheless, it wouldn't be misleading to describe it with-

[27] Why doesn't supervaluationism satisfy (i)? Because the argument in its favour is that it provides an account of vagueness that can accommodate the data. The apparent validity of classical reasoning is part of the data. What's needed for (i) is a direct argument for non-classical logic from an account of the nature of vagueness. The supervaluationist hasn't provided this.

out qualification as either red or orange. For how could b fail to be red or orange? It's clearly coloured. And it's clearly no colour that's neither red nor orange. So surely it's either red or orange, despite being neither clearly red nor clearly orange. This is an instance of a *penumbral connection*: an analytic connection between the borderline regions of vague predicates.[28]

Supervaluationists accommodate penumbral connections by imposing *penumbral constraints*: conditions on the extensions sharpenings assign to penumbrally connected expressions. The intended interpretation(s) of a vague language respects these constraints. For example, the following constraint makes 'b is red $\vee$ b is orange' clearly true, even if both disjuncts are borderline:

> For any sharpening s, either $b \in [\![\text{red}]\!]_s$ or $b \in [\![\text{orange}]\!]_s$, but not both.

Now consider a cube c that's just slightly redder than b, though still borderline red/orange. This should be clearly false:

$$c \text{ is orange} \wedge b \text{ is red}$$

For how could something redder than b fail to be red, if b is? And this should be borderline:

$$c \text{ is red} \wedge b \text{ is red}$$

Yet the conjuncts of both are borderline. The following gives the desired result:

> For any objects x, y and sharpening s, if x is redder than y, then: if $y \in [\![\text{red}]\!]_s$, then $x \in [\![\text{red}]\!]_s$.

This second case also shows that the clear-truth-value of a molecular sentence shouldn't be a function of those of its components. For the conjunctions above differ in clear-truth-value, though their conjuncts don't. Since supervaluationist logic isn't clear-truth-value-functional, this counts in its favour. In fact, the failure of clear-truth-value-functionality arises for any view that preserves classical logic alongside borderline sentences. For when A is borderline, so are the following: A; $A \vee A$; $A \wedge A$; and $\neg A$. But $A \vee \neg A$ will then be clearly true and $A \wedge \neg A$ clearly

[28] This section draws on the argument against clear-truth-value-functionality in Edgington (1997, §3).

false.

We now turn from supervaluationism's benefits to its problems.

2.6 A logical problem

Retention of classical logic was presented as one of supervaluationism's key benefits (§2.5.3). Williamson (1994, §5.3) challenges this. Although his argument isn't quite conclusive, we'll see that it does force surprising revisionary theses on the Supertruth theorist that many will find unacceptable. The Sharpening View, by contrast, survives unscathed.

§2.6.1 presents Williamson's objection. McGee and McLaughlin's response is examined in §2.6.2, alongside two difficulties for it in §§2.6.3–2.6.4. §2.6.5 closes by showing how the Sharpening View evades the objection.

2.6.1 The argument

This section presents the argument against the classicality of supervaluationist logic.

In what sense does the Supertruth View preserve classical logic? Since truth is identified with supertruth, consequence is identified with $\models_{\text{global}}$:

> $\Gamma, A \models_{\text{global}} C$ iff, for any supervaluationist model M, if every member of $\Gamma \cup \{A\}$ is supertrue in M, then C is supertrue in M.

Let $\models_{\text{cl}}$ be classical consequence. Say that an argument is $\models_x$-valid iff its conclusion is an $\models_x$-consequence of the set of its premisses. Say that an inferential pattern is $\models_x$-sound iff it licenses only $\models_x$-valid arguments. $\models_{\text{global}}$ and $\models_{\text{cl}}$ coincide within predicate calculus. All $\models_{\text{cl}}$-valid arguments formalisable within predicate calculus are therefore $\models_{\text{global}}$-valid. Hence any inferential pattern within predicate calculus that is $\models_{\text{cl}}$-sound is also $\models_{\text{global}}$-sound. This is the sense in which the Supertruth View preserves classical logic.

This fails for languages enriched with Δ. Williamson gives counterexamples to *contraposition*, *argument by cases*, *reductio* and *conditional proof*. We'll focus on the last of these:

Conditional Proof (CP) If $\Gamma, A \vdash C$, then: $\Gamma \vdash A \to C$.

'$\Gamma, A \vdash C$' means that C is derivable from the wffs in Γ together with A. Given this rule of proof, the conditional $A \to C$ is derivable from premisses Γ whenever C is derivable from Γ together with A. CP encodes the primary means of drawing conditional conclusions from categorical premisses. CP is $\models_x$-sound iff:

If $\Gamma, A \models_x C$, then: $\Gamma \models_x A \to C$.

CP is $\models_{\text{global}}$-unsound because:

$A \models_{\text{global}} \Delta A$, but: $\not\models_{\text{global}} A \to \Delta A$.[29]

Although $A \models_{\text{global}} \Delta A$, the conditional $A \to \Delta A$ is not a $\models_{\text{global}}$-logical truth. Here's the proof.

First, we show $A \models_{\text{global}} \Delta A$. Suppose A is supertrue in M. Then $s, M \Vdash A$, for all $s \in M$. So by the rule for Δ: $s, M \Vdash \Delta A$, for all $s \in M$. So ΔA is supertrue in M. So if A is supertrue in M, then ΔA is supertrue in M. Since M was arbitrary: $A \models_{\text{global}} \Delta A$. This result shows that the following is $\models_{\text{global}}$-sound:

ΔIn $A \vdash \Delta A$.

We now show $\not\models_{\text{global}} A \to \Delta A$. Suppose that (i) A is not supertrue in M, and (ii) A is not superfalse in M. From (i): $s, M \nVdash A$, for some $s \in M$. By the rule for Δ: $s, M \nVdash \Delta A$, for all $s \in M$. But from (ii): $s, M \Vdash A$, for some $s \in M$. Putting these together: $s, M \Vdash A$ and $s, M \nVdash \Delta A$, for some $s \in M$; let s^* be such a sharpening. By the rule for $\to$: $s^*, M \nVdash A \to \Delta A$. So: $A \to \Delta A$ is not supertrue in M. Hence: $\not\models_{\text{global}} A \to \Delta A$.

The problem comes from combining ΔIn with CP. Suppose A. By ΔIn: ΔA. By CP: $A \to \Delta A$. Since A was our only premiss: $\vdash A \to \Delta A$. But since $\not\models_{\text{global}} A \to \Delta A$: the unpremissed argument for $A \to \Delta A$ is $\models_{\text{global}}$-unsound. Any deductive system containing both ΔIn and CP is therefore $\models_{\text{global}}$-unsound. We could avoid

[29] J.R.G. Williams (2008) argues that this result is an artefact of an impoverished formal setting that does not hold in a more satisfactory framework. There isn't space to discuss Williams's view here. For a response to Williams, see my (forthcoming).

this by excluding ΔIn. But since that rule is sound, this serves only to artificially block derivation of some consequences of our premisses. The Supertruth theorist should therefore deny that CP is sound.

2.6.2 Restricting CP

McGee and McLaughlin (1998, 2004) object. They grant that ΔIn is $\models_{\text{global}}$-sound, but deny that CP *as formulated* is part of classical logic. They claim that a restricted version of CP that disallows ΔIn within the scope of suppositions is all that classical logic requires. Write '$\Gamma, A \vdash_{MM} C$' when C is derivable without using ΔIn from the members of Γ together with A. McGee and McLaughlin claim that classical logic requires not CP, but:

Restricted Conditional Proof (RCP) If $\Gamma, A \vdash_{MM} C$, then: $\Gamma \vdash A \rightarrow C$.

The idea is that although $\models_{\text{cl}}$ and $\models_{\text{global}}$ come apart, this doesn't mandate revisions to classical inferential practice because that practice requires only RCP, not CP. We now have to ask: what is classical logic?

According to Williamson (2004, p.120), classical logic comprises "those forms of logical inference tried and tested in mainstream mathematics and other branches of science." McGee and McLaughlin (2004, p.133) agree. This brings out two virtues of retaining classical logic. The first is that if a classical rule is unsound, then those parts of science and mathematics that employ it become suspect. The second is that an inference rule's successful employment throughout our best science and mathematics provides inductive grounds for believing it successful elsewhere in science, including semantics.

These virtues belong to any semantic theory that renders sound those inferences required by best science and mathematics. McGee and McLaughlin claim that standard predicate calculus without Δ suffices for formalising those inferences. Since $\models_{\text{global}}$ and $\models_{\text{cl}}$ coincide within predicate calculus, the Supertruth View possesses the virtues attendant upon retaining classical logic. Although those relations diverge in languages enriched by Δ, this doesn't bring objectionable revisions to classical inferential practice.

This dispute about the extent of classical logic concerns which of the following dispositions is manifested in mainstream scientific and mathematical reasoning:

> The disposition to conclude $A \to C$ on the basis of any derivation of C from A.
>
> The disposition to conclude $A \to C$ on the basis of any derivation of C from A that doesn't employ ΔIn.

McGee and McLaughlin must claim that only the latter, weaker disposition is manifested in the inferential behaviour of mainstream scientists, if they are to defend RCP over CP. But it is not clear which of these competing accounts is preferable. Attribution of the second disposition is the minimum required to explain the data, if McGee and McLaughlin are right that mainstream science requires only those inferences formalisable in predicate calculus. But then attribution of the first disposition avoids attributing to scientists restrictions on when they are prepared to draw conditional conclusions, when those restrictions aren't manifested in their actual reasoning. It is therefore unclear whether or not the attempted restriction of classical logic to RCP is successful.

2.6.3 The justification for RCP

RCP is *prima façie* objectionable. Why is ΔIn inapplicable to mere suppositions? The rule is sound: the language can't be interpreted so as to make A true without also making ΔA true. Thus ΔA may be inferred from the believed premiss A.[30] So if ΔA cannot be inferred from the mere supposition A, then that must be because the content of A differs between premisses and suppositions: supposing A is not the same as supposing A to be true. This creates two problems. Firstly, an account is required of the content of supposing A, if it isn't the same as supposing A to be true or as believing A. Secondly, it undermines the role of deduction from suppositions in justifying belief in the conclusion of an argument on the basis of (i) belief in its premisses and (ii) a prior deduction of that conclusion from the supposition

[30] Unless believing A and believing A to be true have the same content, then it's obscure what role validity—i.e. *truth*-preservation under every interpretation—has to play in constraining rational belief-formation.

of the premisses. This section finesses McGee and McLaughlin's objection to CP in order to avoid these complaints.

Consider a thinker who supposes A and applies ΔIn to derive ΔA. Granted that her derivation was sound, what exactly has she shown? Since ΔIn is $\models_{\text{global}}$-sound, this derivation shows that any model where A is (super)true is a model where ΔA is also (super)true. It follows that $A \rightarrow \Delta A$ is (super)true in all such models. Does it follow that this conditional is (super)true in *all* models? That's what's needed for $\models_{\text{global}} A \rightarrow \Delta A$, and hence for our thinker's derivation to license the conclusion $A \rightarrow \Delta A$ outside the scope of the initial supposition.

This would follow if the only models where A isn't (super)true were models where A is (super)false. For $A \rightarrow \Delta A$ is vacuously (super)true in any such model. But these aren't the only other models: some models make A neither (super)true nor (super)false. Ensuring that $A \rightarrow \Delta A$ is (super)true in all models where A is (super)true and also (super)true in all models where A is (super)false, therefore doesn't suffice to ensure that A is (super)true in all models. Yet that's all that's ensured by the $\models_{\text{global}}$-validity of our thinker's derivation of ΔA from A. CP fails because in fact, if A is neither supertrue nor superfalse in M, then (i) $s, M \nVdash \Delta A$, for all $s \in M$, and (ii) $s, M \Vdash A$, for some $s \in M$. From (i) and (ii) it follows that $s, M \nVdash A \rightarrow \Delta A$, for some $s \in M$, and hence that $A \rightarrow \Delta A$ is not supertrue in M. This makes M a countermodel to $\models_{\text{global}} A \rightarrow \Delta A$. Hence McGee and McLaughlin's claim that assuming unrestricted CP is tantamount to assuming Bivalence (McGee and McLaughlin, 2004, pp.134–5).

This finesses the restriction on CP. The use of ΔIn within the scope of suppositions is unproblematic. Likewise for drawing conditional conclusions on the basis of such uses of ΔIn. The problem lies in discharging the premiss/supposition to which ΔIn was applied, to yield a conditional without dependence on that premiss/supposition. The $\models_{\text{global}}$-soundness of ΔIn ensures the truth of $A \rightarrow \Delta A$ only under the supposition of A or the supposition of $\neg A$, not under any supposition whatsoever; specifically, not under the supposition $\neg\Delta A \wedge \neg\Delta\neg A$. By discharging the supposition, we lose any record of this information. Hence we cannot do so. No difference in A's content when taken as a premiss or supposition is required because there's no difference in the applicability of ΔIn to premisses and suppositions. The objection to RCP with which this section began therefore fails.

This suggests a sense in which supervaluationist semantics preserves classical reasoning, even within languages enriched by Δ. The countermodels to the classical logical laws can be ignored when reasoning under the (possibly tacit) assumption of precision—i.e. the assumption that there are no borderline cases and Bivalence holds—because those countermodels are all models that make A valueless. Provided we can exclude circumstances in which A is borderline from the circumstances our reasoning must take account of, we can reason classically in a language with supervaluationist semantics. It is certainly plausible that we can exclude borderline cases within pure mathematics. And the replacement of vocabulary susceptible to borderline cases with new classifications is arguably also one of the hallmarks of science. Those, such as mainstream scientists and mathematicians, who reason under the assumption of precision or in circumstances in which borderline cases cannot arise, may therefore reason entirely classically.

2.6.4 A problem for RCP

This section argues that replacing CP with RCP brings unexpected and revisionary consequences.

Given the following pair, RCP's restriction to derivations that don't employ ΔIn is no restriction at all:

(i) Logically valid deductions license the truth of English conditionals, whether subjunctive or indicative.

(ii) English conditionals imply their corresponding material conditionals.

On the Supertruth View, A logically implies ΔA. So by (i): if it were that A, it would be that ΔA. Then by (ii): $A \rightarrow \Delta A$. So any valid argument from A to ΔA also licenses a valid argument for $A \rightarrow \Delta A$ that doesn't employ ΔIn. Hence RCP is equivalent to CP. Those who reject CP but not RCP must therefore reject (i) or (ii). I don't know which is preferable, but neither is attractive and both assumptions are commonplace. Following are three examples.

First example: Ian McFetridge (1990) assumes (i) without argument when arguing that logical necessity is the strongest form of necessity.

Second example: the portion of (i) that concerns counterfactuals follows from the Lewis-Stalnaker semantics. Since the logical consequences of A are true at any world where A is true, those consequences are also true at the closest world(s) where A is true.

Third example: Williamson (2007, pp.293–4, 300) calls the portion of (ii) that concerns counterfactuals "immensely plausible", noting that it is an *axiom* of Lewis's logic for counterfactuals (Lewis, 1986a, p.132).

These appeals to the authority of classical logicians don't show that (i) and (ii) are true. But they do expose the Supertruth View's (well hidden) revisionary implications.

2.6.5 The Sharpening View

The Sharpening view renders CP sound. On that view, consequence is local consequence:

> $\Gamma, A \models_{\text{local}} C$ iff, for any model M and sharpening $s \in M$, if $s, M \Vdash \Gamma \cup \{A\}$, then $s, M \Vdash C$.

ΔIn is $\models_{\text{local}}$-unsound. For suppose that (i) A is not supertrue in M, and (ii) A is not superfalse in M. By (i) and the rule for Δ: $s, M \nVdash \Delta A$, for any $s \in M$. By (ii): $s, M \Vdash A$, for some $s \in M$. Putting these together: $s, M \Vdash A$ and $s, M \nVdash \Delta A$, for some $s \in M$. Hence: $A \not\models_{\text{local}} \Delta A$.

In fact, the Sharpening View's formal treatment of truth, consequence and Δ is exactly analogous to that of truth, consequence and $\Box$ in standard possible-worlds semantics for modal logic. On the Sharpening View, vagueness mandates no more nor less deviation from classical logic than does modality.

2.6.6 Supervaluationist logic: concluding remarks

We've seen two kinds of problems for the Supertruth theorist's claim to preserve classical logic. The first is that it's unclear whether mainstream scientific reasoning manifests disposition to reason in accord with the restricted or unrestricted version of CP (§2.6.2). This first problem is somewhat ameliorated by the fact that unrestricted classical reasoning is permissible when, as in most science and mathematics, the possibility of borderline cases may be discounted (§2.6.3). The second is that it brings revisionary consequences for the interaction of logical consequence with English and material conditionals (§2.6.4). Although neither problem is decisive, they are costs of the Supertruth View that aren't incurred by the Sharpening View (§2.6.5).

2.7 Two semantic problems

Let us turn from supervaluationist logic and inference to the semantics on which they are based. §2.7.1 examines the Supertruth theorist's notion of truth. §2.7.2 presents a difficulty in accommodating our apparent discretion about borderline classification. The Supertruth View will again be shown to fare significantly less well than the Sharpening View.

2.7.1 Truth and supertruth

Because borderline sentences are neither (super)true nor (super)false, the Supertruth View violates:

Bivalence For any sentence A, either A is true or A is false.

But since sharpenings are just classical models, every classical theorem is true at each sharpening and hence supertrue in each model. The following is therefore sound:

Law of Excluded Middle (LEM) Every instance of $\ulcorner A \vee \neg A \urcorner$ is a theorem.

LEM is one component of the Supertruth theorist's claim to preserve classical logic. This section shows that this combination of views makes the identification of truth with supertruth problematic.

2.7.1.1 Supertruth and Bivalence

The Supertruth View implies that A and $\ulcorner A$ is true$\urcorner$ differ in content. $\ulcorner A$ is true$\urcorner$ should be true at a sharpening iff A is supertrue:

$$s, M \Vdash \ulcorner A \text{ is true}\urcorner \text{ iff } A \text{ is supertrue in } M; \text{ iff } t, M \Vdash A, \text{ for all } t \in M.$$[31,32]

Suppose A is neither supertrue nor superfalse in M. So $s, M \nVdash A$, for some $s \in M$. By the rule for 'is true': $s, M \nVdash \ulcorner A$ is true$\urcorner$, for all $s \in M$. So $\ulcorner A$ is true$\urcorner$ is (super)false in M. But since A is not (super)false in M but valueless, A and $\ulcorner A$ is true$\urcorner$ differ in (super)truth-value and therefore also differ in content.

Recall the Supertruth View's account of clear truth as truth and clear falsity as falsity (§2.5.1.1). This account of clarity is incomplete without an account of the difference in content between A and $\ulcorner A$ is true$\urcorner$. The Supertruth View's own account of clarity is the source of this extra explanatory burden. For that account combines with the possibility of borderline cases to undermine Bivalence; and non-Bivalence is responsible for the difference in truth-status of A and $\ulcorner A$ is true$\urcorner$ when A is borderline. Without an account of the difference in content between A and $\ulcorner A$ is true$\urcorner$, the Supertruth View's explanation of clarity in semantic terms is only a pseudo-explanation.

Let us be clear about just what this shows. It does not show that the truth of $\ulcorner A$ is true$\urcorner$ requires anything more than the truth of A, or *vice versa*; for the models that make A true are exactly those that make $\ulcorner A$ is true$\urcorner$ true. The difference in content comes from models that make A valueless and $\ulcorner A$ is true$\urcorner$ false. Although the truth of $\ulcorner A$ is true$\urcorner$ requires no more nor less than the truth of A, its falsity requires less than the falsity of A. This is what needs explaining.

2.7.1.2 From LEM to Bivalence

What account of this difference in content between A and $\ulcorner A$ is true$\urcorner$ might the Supertruth theorist offer? This section examines their conception of sentence-content, arguing that it is unclear how such an account might proceed.

Classical semantics conceives sentence-content as comprising a truth-condition.

[31] The alternative is: $s, M \Vdash \ulcorner A$ is true$\urcorner$ iff $s, M \Vdash A$. But then each instance of Bivalence is true at each sharpening: A is true $\vee\neg(A$ is true). This either (i) reinstates Bivalence, or (ii) prevents the object-language 'is true' from expressing genuine (super)truth.

[32] We use a truth-predicate and ignore the semantic paradoxes for simplicity. We could could just as easily use a truth-operator and forego reference to sentences, in which case the paradoxes couldn't arise.

If the condition is met, the sentence is true, otherwise it is false. The Supertruth View cannot accept this. For if truth-conditions exhaust sentence-content, then there is no space to distinguish the untruths that are false from those that fall down a truth-value gap: in both cases, the condition that exhausts the content of the sentence in question is unsatisfied. An alternative conception of sentence-content is required.[33]

The Supertruth theorist needs to conceive sentence-content as comprising two independently settled components: truth-conditions and falsity-conditions. A sentence is true iff its truth-condition is met, false iff its falsity-condition is met, and neither true nor false when neither condition is met. When the content-determining facts co-operate, truth-conditions and falsity-conditions will partition the possibilities and the (interpreted) sentence in question will be Bivalent. But the meaning-determining facts need not co-operate; in which case, truth- and falsity-conditions won't partition the possibilities and the sentence won't be Bivalent; in some possibilities, it will be neither true nor false because neither its truth-conditions nor falsity-conditions are satisfied.

We can adapt an argument of Williamson's (1997, §1) to make trouble for this view. Since truth just is the satisfaction of truth-conditions, and falsity just is the satisfaction of falsity-conditions, the following are analytic:

TC *A*'s truth-condition is satisfied iff *A* is true.

FC *A*'s falsity-condition is satisfied iff *A* is false.

Take 'grass is green' as an example. A plausible truth-condition is that grass is green. And a plausible falsity-condition is that grass is not green. By LEM: grass is green $\vee$ grass is not green. Suppose that grass is green. By TC: 'grass is green' is true. By $\vee$-introduction: 'grass is green' is true $\vee$ 'grass is green' is false. Now suppose that grass is not green. By FC: 'grass is green' is false. By $\vee$-introduction:

[33] What about (putative) truth-value gaps resulting from non-referring singular terms? The route from reference-failure to valuelessness conceives of singular terms as contributing their referent to truth-conditions. When no referent is contributed, no truth-condition is determined. Hence sentences containing non-referring terms lack truth-conditional content. The Supertruth theorist, by contrast, conceives gappy vague sentences as having a gappy content.

'grass is green' is true $\vee$ 'grass is green' is false. So either way: 'grass is green' is true $\vee$ 'grass is green' is false. Since 'grass is green' was arbitrary, we can generalise to Bivalence. Since TC and FC are analytic and we made no non-analytic assumptions: Bivalence is analytic.

The Supertruth theorist must resist. How? Not by attacking the reasoning. That requires only argument by cases, universal generalisation, *modus ponens* and $\vee$-introduction. The last three of these are $\models_{\text{global}}$-valid. Argument by cases is trickier. Although $\models_{\text{global}}$-valid within predicate calculus, the following result shows that it fails in languages containing Δ:

$\models_{\text{global}} A \vee \neg A.$

$A \models_{\text{global}} \Delta A.$

$\neg A \models_{\text{global}} \Delta\neg A.$

$\not\models_{\text{global}} \Delta A \vee \Delta\neg A.$

On the Supertruth View, Δ amounts to an object-language reflection of (super)truth. This result might therefore appear to cast doubt on arguing by cases from TC, FC and LEM to Bivalence. This appearance is misleading.

$A \models_{\text{global}} \Delta A$ and $\neg A \models_{\text{global}} \Delta\neg A$ together ensure that one of ΔA, $\Delta\neg A$ will be supertrue in any model where one of A, $\neg A$ is supertrue. Hence $\Delta A \vee \Delta\neg A$ will also be supertrue in any such model. But this is silent about models where neither A nor $\neg A$ is supertrue. Since $A \vee \neg A$ is supertrue in some models where neither disjunct is supertrue, $A \models_{\text{global}} \Delta A$ and $\neg A \models_{\text{global}} \Delta\neg A$ do not by themselves ensure that $\Delta A \vee \Delta\neg A$ is supertrue in all models where $A \vee \neg A$ is supertrue. Hence even given $\models_{\text{global}} A \vee \neg A$, they do not ensure that $\Delta A \vee \Delta\neg A$ is supertrue in all models whatsoever. The countermodels to the validity of arguing by cases arise because both ΔA and $\Delta\neg A$ are in fact false at all sharpenings in all models where neither A nor $\neg A$ is supertrue. Nothing similar applies to the argument by cases from TC and FC to Bivalence.

On the Supertruth View's bipartite conception of sentence-content, the contents of $\ulcorner A$ is true$\urcorner$ and of $\ulcorner A$ is false$\urcorner$ are exhausted by A's truth-conditions and falsity-conditions respectively. Thus TC and FC should be understood as analytically true material biconditionals, not mutual entailments.[34] Since TC and FC are analytic, they should be supertrue in all models. So if, in M, B expresses A's truth-condition, then B and $\ulcorner A$ is true$\urcorner$ have the same truth-value at all sharpenings in M because they have the same content. And if, in M, C expresses A's falsity-condition, then C and $\ulcorner A$ is false$\urcorner$ have the same truth-value at all sharpenings in M because they have the same content. So if the truth-condition for 'grass is green' is that grass is green, and if the falsity-condition for 'grass is green' is that grass is not green, then any sharpening where either 'grass is green' is true or 'grass is not green' is true is a sharpening where ''grass is green' is true $\vee$ 'grass is green' is false' is true. By LEM: every sharpening is one where either 'grass is green' is true or 'grass is not green' is true. So every sharpening is one where ''grass is green' is true $\vee$ 'grass is green' is false' is true. So 'grass is green' is Bivalent. The argument by cases from TC and FC to Bivalence is therefore $\models_{\text{global}}$-valid. The Supertruth theorist must resist its premisses.

The only premisses were (i) TC, (ii) FC, (iii) the truth-condition for 'grass is green' is that grass is green, and (iv) the falsity-condition for 'grass is green' is that grass is not green. Since TC and FC are components of the Supertruth theorist's account of sentence-content, they must reject (iii) or (iv). Symmetry suggests they will reject both. But what replacements will they offer? Since the analysis of clarity in terms of truth is the source of this commitment, that analysis is incomplete (since it has barely even begun) until replacements are supplied. No Supertruth theorist has yet done so.

Given the fundamental nature of truth, it is *prima façie* doubtful whether the Supertruth theorist can supply alternatives to 'grass is green' and 'grass is not green' as truth- and falsity-conditions for 'grass is green'. Thus it is doubtful

[34] The Supertruth theorist denies that mutual entailment is sufficient for sharing of truth-conditional content. Only a $\models_{\text{global}}$-valid material biconditional suffices for identity of truth-value at all sharpenings. Since the Supertruth View cashes out content via distribution of truth-value across a space of sharpenings, only a $\models_{\text{global}}$-valid material biconditional suffices for identity of content. Thus if A expresses the truth-condition of B, then $\models_{\text{global}} A \leftrightarrow (B \text{ is true})$, provided we restrict models to those that respect either the intended senses of A and B, or the intended relationship between their senses; including an axiom for 'is true' as a base clause in the recursive definition of $\Vdash$ enforces this second restriction.

whether the Supertruth theorist can explain their conception of truth. Should we expect them to be able do so? The demand for an account of the truth-conditions of 'grass is green' is a demand for a sentence whose content is exhausted by one of the two conditions that together comprise the content of 'grass is green'. How might such a sentence enter our language? Not via the same route as 'grass is green'; for then the two would be vague in just the same ways. The most we have any right to expect, the Supertruth theorist may claim, is an infinite disjunction, each of whose disjuncts describes a possible state sufficient for grass to be green (and there seems little reason to expect even that). Unfortunately, no such sentence will be statable. In order to make it statable, we need a condition ϕ common to exactly those states in which one of the disjuncts is satisfied—that is, a condition common to exactly those states sufficient for grass to be green—so that we can say that one of them obtains:

'Grass is green' is true iff $\exists x(\phi x)$.

Yet the problem remains: what right have we to expect an English sentence other than ''grass is green' is true' or 'grass is green' that will be true in exactly those possibilities where grass is green? And what right have we to expect an English predicate true of exactly those states sufficient for grass to be green, other than 'is a state of grass's being green'? Yet without such expressions, any account of ϕ will be unilluminating.

An inability to state an adequate truth-condition for 'grass is green' other than ''grass is green' is true' is insufficient to refute the Supertruth theorist because they are not committed to there being any informative account of such a condition. But it should make us reluctant to endorse the view. For until an appropriate condition has been informatively specified, we lack guarantee that the Supertruth theorist's conception of truth is both contentful and coherent. Since the Supertruth View explains clarity in terms of truth, we lack guarantee, or even positive reason to believe, that the proposed explanation of clarity is both contentful and coherent. If truth is super-truth, then non-standard accounts of truth and falsity are required. These have not been provided and it is doubtful that they could be. The Supertruth View thus incurs a possibly un-meetable explanatory obligation.

2.7.1.3 The Sharpening View

This problem does not afflict the Sharpening View because on that view, truth under an interpretation is truth under a classical interpretation. For any sentence A, model M and sharpening $s \in M$, exactly one of the following holds:

$s, M \Vdash A$.

$s, M \Vdash \neg A$.

Since A is false iff $\neg A$ is true, the Sharpening view satisfies Bivalence. The following axiom for 'is true' ensures that A and $\ulcorner A$ is true$\urcorner$ receive the same truth-value at every sharpening and therefore express the same content under every interpretation that respects the intended sense of 'is true':

$s, M \Vdash \ulcorner A$ is true$\urcorner$ iff $s, M \Vdash A$.

Because the Sharpening theorist doesn't treat clear truth as a semantic classification, but as a partly semantic and partly metasemantic classification, they can endorse any account of truth available to the classical semanticist.

2.7.2 Borderline discretion

§2.3.3.2 observed that we must occasionally decide whether to count a borderline case as a positive case or a negative case. This commonplace feature of linguistic usage brings none of the discomfort of misuse: it seems compatible with (and perhaps even partially constitutive of) competence with the expressions in question. This section argues that the Supertruth View cannot accommodate this.

2.7.2.1 Supertrtuth and borderline discretion

Let a be a borderline F. Suppose you are in a situation where a decision is required about whether or not to count a as an F. Either choice is open to you. Because such situations are commonplace and unremarkable, a maximally satisfactory semantics would allow for either decision without misclassification. The Supertruth View cannot do so.

On the Supertruth View, neither Fa nor $\neg Fa$ is true when a is a borderline F. So to count a as an F or to count it as a non-F, is to misclassify it. If truth is supertruth, then this feature of our use of vague language is, strictly, a misuse of language. But since meaning is determined by use, an expression's semantic properties should be compatible with most aspects of its use, and certainly with all of its most deeply entrenched ones. Indeed, the legitimacy of those uses should flow naturally from the correct semantic theory. If truth is super-truth, then this is not so. The truth-value gaps used to explain borderline ignorance provide too strong an explanation, one that renders seemingly legitimate uses of language illegitimate.

2.7.2.2 Sharpenings and borderline discretion

The Sharpening View lessens the problem without dissolving it entirely. On that view, borderline status is not a semantic status incompatible with truth and incompatible with falsity. Instead, the sharpenings provide a range of semantic classifications, only some of which are incompatible with counting a borderline F as, say, an F. Doing so still involves misclassification, but it also involves correct classification. If we can treat these as cancelling each other out, then a perfectly balanced borderline case can be counted either way without misclassification (though without correct classification also).

The Sharpening View's core thesis is that the meaning-determining facts determine many intended interpretations. We've just seen that this weakens the objection from borderline discretion. A natural addition to the view eliminates it entirely. This addition allows decisions about the classification of borderline cases to count amongst the meaning-determining facts in such a way that deciding to count a as, say, an F suffices for interpretations that make Fa false to count as unintended, provided a is a borderline F: classificatory decisions about borderline Fs settle their status with regard to F by narrowing the semantic properties of F.

This approach makes it context-sensitive just which interpretations are intended. The notion of an intended interpretation ought therefore to be relativised to a particular conversational context or sub-community of a whole linguistic community; for otherwise my decision to count a borderline F as an F would affect the legitimacy of your decision not to, even if you are in a different town from me. This relativisation is natural if we think of intended interpretations as encoding the information communicated by uses of language: those privy to my decision to count a as an F can recover different information from my uses of Fa than those not privy to that decision. On this view, the intended interpretations of my and my listener's shared language vary across contexts, depending on whether I am communicating with one group or the other because different groups can recover different information from my utterances.

These temporary classificatory decisions shouldn't affect which things count as borderline cases: deciding to count a terracotta pot as a red pot doesn't prevent it from being borderline red/orange. There are two natural and complementary ways of achieving this. According to the first, the whole community's language use determines a range of intended interpretations. These settle the borderline cases and limit the classificatory decisions available to the community's members. The second approach relativises the notion of a borderline case to a linguistic community or context. A community's use of language settles a range of intended interpretations that limit the classificatory decisions available to the members of its sub-communities. The decisions of these sub-communities c don't affect the intended interpretations determined by the linguistic behaviour of the wider community c^*, and hence don't affect what counts as borderline relative to c^* despite affecting what counts as borderline relative to c. Whichever approach we prefer, the borderline cases of English predicates are invariant across the community of English speakers, despite the decisions of particular speakers affecting the intended interpretations of their utterances within the contexts in which those classificatory decisions are made.

Is a similar response available to the Supertruth theorist? Can't they also treat the class of sharpenings as context-sensitive and responsive to our classificatory decisions about borderline cases? Maybe they can. But there is a difficulty to overcome first. Context-sensitivity of sharpenings can make it correct to count a borderline F as an F following a decision to do so. This does not however, legitimise making that initial decision; for that decision was made in a context where the borderline F in question occupied a semantic status incompatible with its being an

F. Decisions to count borderline *F*s one way or the other are decisions to count it as something it is not; they are decisions to mis-classify. The problem for the Supertruth View isn't whether we can ultimately judge borderline classificatory decisions correct once they have been made, but whether we can legitimately make them in the first place. Nothing similar affects the Sharpening View.

2.7.3 Supervaluationist semantics: concluding remarks

We've seen two problems for supervaluationist semantics. One concerned the identification of truth with supertruth. We saw that this incurs a possibly unsatisfiable explanatory burden. This doesn't refute the Supertruth View, but it does (i) create doubt about whether that view is both contentful and coherent, and (ii) undermine the Supertruth theorist's claim to offer an informative analysis of clarity. This problem does not afflict the Sharpening View.

The second problem was that the Supertruth View seems unable to accommodate the legitimacy of temporary decisions about the classification of borderline cases. Although such decisions are commonplace, unremarkable and practically indispensable, the Supertruth View regards them as misuses of the expressions in question, given their semantic properties. The difficulty was shown to be less pressing for the Sharpening View.

The following two sections present two further problems for the Supertruth View.

2.8 Field on truth and super-truth

It seems misguided even to try and discover when was the last second of Bertrand Russell's youth.[35] Hartry Field objects to the Supertruth theorist's explanation we presented in §2.5.1.1:

> "The supervaluationist says that at certain stages, Russell was neither in the determinate positive extension nor the determinate negative extension of 'old'. But of what possible interest is this, given that (according

[35] Suppose for simplicity that time isn't dense.

> to the view) he was at those stages either old or not old?" (Field, 2008, p.155)

The Supertruth theorist has a reply. If truth is supertruth, then a predicate's determinate positive and negative extensions are its positive and negative extensions: the sets of things of which it is true and of which it is false, respectively. So when Russell was borderline old, neither 'Russell is old' nor 'Russell is not old' was true. Since knowledge implies truth, it is neither knowable that Russell was then old, nor knowable that he was not.

Field responds:

> "[C]onsider the question of why a sentence being indeterminate precludes our knowing it. Calling indeterminateness "lack of truth value" might appear to provide an answer: you can't know what isn't true, and if indeterminate sentences lack truth value then you obviously can't know them! But this is just more verbal hocus pocus: what underlies the claim that you can't know what isn't true is that you can't know that p unless p. You can't know that Russell was old at n nanoseconds unless he was old at n nanoseconds, and you can't know that he wasn't old at n nanoseconds unless he wasn't old at n nanoseconds. But on the supervaluationist view he either was or wasn't, and if you can't know which, that needs an explanation. The use of 'true' to mean super-true just serves to disguise this." (Field, 2008, p.154)

This shouldn't immediately convince the Supertruth theorist. For she claims that when Russell was borderline old, the disjunction 'Russell is old or Russell isn't old' was true, but neither disjunct was: it wasn't the case that Russell was old, and it wasn't the case that Russell wasn't old. This is contradictory if 'it's not the case that Russell is not old' involves two occurrences of the same kind of negation: $\neg\neg A$. The Supertruth theorist therefore needs the outer negation to form a truth from any untruth and the inner one to form a truth only from falsehoods (and form valueless sentences from other valueless ones). This casts doubt on the inner (strong) negation sign's claim to express genuine negation. The outer (weak) negation is what's used to explain why we can't know whether Russell was old when he was borderline old; only in that sense is it not the case that A and not the case that $\neg A$ when A is valueless. But only the inner (strong) negation is governed by supervaluationist semantics; for on that semantics $\neg A$ and $\neg\neg A$ are both valueless when A is borderline. The Supertruth View either gives the wrong account of negation or cannot explain borderline ignorance.

This doesn't touch the Sharpening View. §2.5.1.1 offered two candidate necessary conditions on the truth of $\ulcorner S$ knows that $A\urcorner$ under an intended interpretation s:

> The proposition s assigns to A is true.
>
> Each proposition assigned to A by any intended interpretation is true.

Both ensure that $\ulcorner S$ knows that $A\urcorner$ is no better than borderline when A is borderline. Given the following rule, it follows that we ought not claim to know that A when A is borderline:

> Assert only the truth.

Borderline status thus makes investigation into known borderline claims misguided by making it in-principle illegitimate to claim to know the result of the investigation. Since this doesn't appeal to the untruth of borderline claims, the Sharpening View is immune to Field's objection.

2.9 Higher-order vagueness

We introduced vagueness as the fuzziness characteristic of the red/orange, tall/not tall and intelligent/unintelligent distinctions. The extent of this fuzziness is itself fuzzy. This gives rise to the phenomenon of higher-order vagueness. This section examines some objections to supervaluationist accounts of it.

2.9.1 Terminology

We begin with some terminology. Williamson (1999) develops these ideas more carefully.

Consider the classification of objects into the Fs and the non-Fs. Call this the zero-order F-classification. Objects may clearly belong to one of its sub-classifications. F is first-order vague iff there could be objects that don't clearly belong to any sub-classification of the zero-order F-classification; iff there could be borderline cases to the zero-order F-classification. Such first-order borderline cases of F are neither clearly F nor clearly not F.

Consider this last classification into the clear Fs, first-order borderline Fs and clear non-Fs. Call this the first-order F-classification. Objects may clearly belong to one of its sub-classifications. F is second-order vague iff there could be objects that don't clearly belong to any sub-classification of the first-order F-classification; iff there could be borderline cases to the first-order F-classification. Such second-order borderline cases of F are either:

(i) neither clearly clearly F nor clearly not clearly F; or

(ii) neither clearly first-order borderline F nor clearly not first-order borderline F; or

(iii) neither clearly clearly not F nor clearly not clearly not F.

Consider this last classification into the (i) clearly clear Fs, (ii) borderline cases of clear Fs, (iii) clearly first-order borderline Fs, (iv) borderline cases of first-order borderline Fs, (v) clearly clearly not Fs, and (vi) borderline cases of clearly not Fs. Call this the second-order F-classification. Objects may clearly belong to one of its sub-classifications (i)–(vi). F is third-order vague iff there could be objects that don't clearly belong to any sub-classification of the second-order F-classification; iff there could be borderline cases to the second-order F-classification. We won't list the possibilities for these third-order borderline Fs.

Iterating this construction allows us to define arbitrarily high orders of borderline case and vagueness. We can extend it to borderline *sentences* via the stipulation that an ith-order borderline sentence is an ith-order borderline case of a truth. F is precise iff F is not ith-order vague, for any $i > 0$. F is higher-order vague iff F is ith-order vague for some $i > 1$.

2.9.2 Varieties of higher-order vagueness

This section examines the relationship between the technical notions defined in the previous section and the phenomenon they are intended to capture.

The arguments for first-order vagueness extend naturally to higher-order vagueness. There seems no non-arbitrary stopping point. It might be objected that the world itself may not be fine-grained enough to allow distinctions between every definable order of borderline case, and hence that, above some level, the orders collapse into one. This may be right. But this collapse is imposed by the world our vague concepts describe, rather than by those concepts themselves. An adequate analysis of vagueness ought not to presuppose it.

Another objection to arbitrarily high orders of vagueness is that it rapidly outstrips our capacity to comprehend; understanding attributions of third-order vagueness is, for most, a very difficult task. But difficulties with understanding don't imply non-existence. The case is similar to arbitrarily long and complex sentences of English. They may not be comprehensible, but they are still meaningful. Roy Sorensen's (2010) contains related discussion and arguments for arbitrarily high orders of vagueness.

These considerations motivate the thesis of:

Unrestricted Borderline Cases (UBC) If F is vague, then there could be borderline cases to any sub-classification of the ith-order F-classification, for all i. There are no limits on higher-order vagueness.[36]

Compatibility with UBC is an attractive feature of theories of vagueness. Note however that this technical notion is intended to capture an intuitive idea of ineradicable fuzziness:

No Sharp Boundaries (NSB) Vague predicates impose only fuzzy classifications, they mark no sharp boundaries whatsoever.

What is the relationship between UBC and NSB? There are two attitudes one might take.

[36] UBC is stronger than the following: vague predicates are ith-order vague, for all i. UBC requires borderline cases to every sub-classification of the ith-order classification, while this weaker alternative only requires borderline cases to some sub-classification of the ith-order classification.

The first approach takes UBC as an analysis of NSB. This view identifies fuzziness with the (possible) presence of borderline cases. The ineradicable fuzziness of NSB is then identified with there being no restrictions on higher-order vagueness.

The alternative view takes UBC as a consequence of NSB, but not as an analysis. On this view, ineradicable fuzziness is responsible for the existence of unrestrictedly high orders of borderline case, though the content of the former notion may outstrip that of the latter. We might even add that the fuzziness of the initial F/non-F classification is what's responsible for NSB: to be fuzzy is to be ineradicably fuzzy, and hence to entirely lack sharp classificatory boundaries. On this view, the orders of vagueness are manifestations of the underlying phenomenon of fuzziness.

It shouldn't matter to our discussion which approach is correct. We'll focus on UBC. Since it's very hard to see how a classification could be fuzzy and yet not allow for borderline cases, UBC seems necessary for NSB. Subsequent sections examine two arguments against UBC.

2.9.3 Higher-order vagueness and the Supertruth View

On the Supertruth View, if A is first-order borderline, then it lacks truth-value. What about if A is second-order borderline? We have three options: true, valueless, and false. The second-order borderline cases therefore collapse into the same semantic status as either the clear cases, the first-order borderline cases or the clear non-cases. Higher-order vagueness is not distinctive at the level of truth-evaluation.

The Supertruth theorist might respond by introducing more truth-values. This brings three problems: (i) it complicates the theory; (ii) the complication is severe because UBC implies that there are infinitely many orders of borderline case, and hence also truth-values; (iii) each extra truth-value requires philosophical explanation.[37] An alternative would be preferable.

This suggests that if truth is super-truth, then higher-order vagueness in F is not a feature of the original F/non-F classification:

[37] This explanatory burden is new. Our Supertruth View employs truth-value *gaps* in place of a third *value*.

"It may be misleading to think of higher-order vagueness in α as a *species* of vagueness in α. Higher-order vagueness in α is first-order vagueness in certain sentences containing α." (Williamson, 1999, p.140)

Focus on the clear end of a Sorites series. The second-order borderline sentences have the same truth-status as either the clear or the first-order borderline sentences. The difference emerges in the truth-status of sentences containing them: ΔA is true when A is clearly true, false when A is first-order borderline, and valueless when A is second-order borderline. Let's accommodate this formally.

2.9.4 Semantics for higher-order vagueness

We want to complicate the supervaluationist formalism to allow distinctions amongst varieties of borderline case. We begin with second-order borderline cases, then third-order borderline cases, and then arbitrarily high-ordered borderline cases. Our strategy iterates the supervaluationist construction to allow distinctions within a model-structure amongst the sentences that fall down a truth-value gap.

Remove Δ from the object-language. We'll replace it with something more adequate shortly. Supervaluationist models are re-named 1-models. A *2-model* M_2 is a class of 1-models. Our original base clauses are amended with an additional relativisation of $\Vdash$ to 2-models:

$v, s, M_1, M_2 \Vdash \Phi^n \alpha_1, \ldots, \alpha_n$ iff $\langle [\![\alpha_1]\!]_{s,v}, \ldots, [\![\alpha_n]\!]_{s,v} \rangle \in [\![\Phi^n]\!]_s$.

$v, s, M_1, M_2 \Vdash \neg A$ iff $v, s, M_1, M_2 \nVdash A$.

$v, s, M_1, M_2 \Vdash A \wedge B$ iff $v, s, M_1, M_2 \Vdash A$ and $v, s, M_1, M_2 \Vdash B$.

$v, s, M_1, M_2 \Vdash \forall x A$ iff $v', s, M_1, M_2 \Vdash A$, for every assignment v' that differs from v at most over 'x'.

$s, M_1, M_2 \Vdash A$ iff $v, s, M_1, M_2 \Vdash A$, for all assignments v.

Define supertruth and superfalsity in a 2-model:

A is supertrue in M_2 iff $s, M_1, M_2 \Vdash A$, for all 1-models $M_1 \in M_2$ and sharpenings $s \in M_1$.

A is superfalse in M_2 iff $s, M_1, M_2 \nVdash A$, for all 1-models $M_1 \in M_2$ and sharpenings $s \in M_1$.

Supertruth (superfalsity) in a 2-model is supertruth (superfalsity) in all of its 1-models. The Supertruth View now identifies truth under an interpretation with supertruth in a 2-model, and falsity under an interpretation with superfalsity in a 2-model. Plugging this into the Tarskian analysis of consequence gives a new account of global consequence:

$\Gamma \models_{\text{global}} C$ iff, for every 2-model M_2, if every member of Γ is supertrue in M_2, then C is supertrue in M_2.

To express claims about first-order vagueness, we add a sentential operator Δ^1:

$v, s, M_1, M_2 \Vdash \Delta^1 A$ iff $v, t, M_1, M_2 \Vdash A$, for all sharpenings $t \in M_1$.

Then we add another operator Δ^2 for expressing claims about second-order vagueness:

$v, s, M_1, M_2 \Vdash \Delta^2 A$ iff $v, s, N_1, M_2 \Vdash A$, for all 1-models $N_1 \in M_2$.

With Δ^1 and Δ^2 in place, first-order borderline cases are distinguishable from second-order borderline cases.

Third-order borderline cases are accommodated by a further iteration. A 3-*model* is a class of 2-models. Further relativise the base clauses to 3-models. Define supertruth and superfalsity in a 3-model:

A is supertrue in M_3 iff $s, M_1, M_2, M_3 \Vdash A$, for all 2-models $M_2 \in M_3$, 1-models $M_1 \in M_2$ and sharpenings $s \in M_1$.

A is superfalse in M_3 iff $s, M_1, M_2, M_3 \nVdash A$, for all 2-models $M_2 \in M_3$, 1-models $M_1 \in M_2$ and sharpenings $s \in M_1$.

Truth and falsity under an interpretation are then identified with supertruth and superfalsity in a 3-model. Consequence becomes supertruth-preservation in every 3-model. Finally, a Δ^3 operator is introduced for expressing claims about third-order borderline cases:

$v, s, M_1, M_2, M_3 \Vdash \Delta^3 A$ iff $v, s, M_1, N_2, M_3 \Vdash A$, for all 2-models $N_2 \in M_3$.

And so on upwards. The construction can be iterated indefinitely to capture indefinitely high orders of vagueness. The general forms of the rules for supertruth and superfalsity in an i-model, and for each Δ^i operator are:

> A is supertrue in M_i iff $s, M_1, \ldots, M_i \Vdash A$, for all $(i-1)$-models $M_{i-1} \in M_i$, $(i-2)$-models $M_{i-2} \in M_{i-1}, \ldots$, and sharpenings $s \in M_1$.
>
> A is superfalse in M_i iff $s, M_1, \ldots, M_i \nVdash A$, for all $(i-1)$-models $M_{i-1} \in M_i$, $(i-2)$-models $M_{i-2} \in M_{i-1}, \ldots$, and sharpenings $s \in M_1$.
>
> $v, s, M_1, \ldots, M_i \Vdash \Delta^i A$ iff $v, s, M_1, \ldots, M_{i-1}, M_i \Vdash A$, for all $(i-1)$-models $M_{i-1} \in M_i$.

Two closing comments. Firstly, if $\Delta^i A$ is supertrue in an i-model, then so is $\Delta^{i-1} A$, as it should be. Secondly, falsity at any sharpening suffices for untruth at any 1-model containing it; which suffices for untruth at any 2-model containing that 1-model; which suffices for untruth at any 3-model containing that 2-model; which suffices.... The merest hint of unclarity suffices for untruth. Different orders of borderline case are distinguished not by their semantic relationship to F, but to open sentences containing F, e.g.: $\Delta^1 Fx$.

2.9.5 Hidden sharp boundaries?

The following are supertrue in any 1-model:

(S4) $$\Delta^1 A \to \Delta^1 \Delta^1 A$$

(S5) $$\neg \Delta^1 A \to \Delta^1 \neg \Delta^1 A$$

It might seem that this rules out higher-order vagueness: neither the clear cases nor the less-than-clear cases can have borderline cases. But Δ^1 is only intended to express claims about first-order borderline cases, not higher-order ones. Iteration of Δ^1 is an artefact of the formation rules without representational import.

It would be bad news if the following were supertrue in any i-model:

(S4i) $$\Delta^{i-1} A \to \Delta^i \Delta^{i-1} A$$

(S5i) $$\neg \Delta^{i-1} A \to \Delta^i \neg \Delta^{i-1} A$$

(S4i) rules out borderline cases to the Δ^{i-1} cases. (S5i) rules out borderline cases to the less than Δ^{i-1} cases. Were either supertrue on all i-models, such kinds of borderline case would be logically impossible and UBC would be false. Fortunately, both fail. Let M_2 be a 2-model containing only the 1-models M_1, N_1 such that:

(2) $s, M_1, M_2 \Vdash A$, for all $s \in M_1$

and:

(3) $s, N_1, M_2 \nVdash A$, for all $s \in N_1$

We now show that neither (S4i) nor (S5i) is supertrue in M_2 (for $i = 2$). Variables ranging over sharpenings are treated as implicitly universally quantified to help with presentation.

From (3) and the rule for Δ^1: $s, N_1, M_2 \nVdash \Delta^1 A$. So by the rule for Δ^2: $s, M_1, M_2 \nVdash \Delta^2\Delta^1 A$. But from (2) and the rule for Δ^1: $s, M_1, M_2 \Vdash \Delta^1 A$. Instantiating (S4$i$) for $i = 2$ therefore gives a conditional $\Delta^1 A \rightarrow \Delta^2\Delta^1 A$ whose antecedent is true at s, M_1, M_2 and whose consequent is not. Hence: $s, M_1, M_2 \nVdash \Delta^1 A \rightarrow \Delta^2\Delta^1 A$. So (S4$i$) is not supertrue in M_2 (for $i = 2$).[38] The argument generalises to show that for no i is (S4i) supertrue on all i-models.

From (2) and the rules for Δ^1 and $\neg$: $s, M_1, M_2 \nVdash \neg\Delta^1 A$. So by the rule for Δ^2: $s, N_1, M_2 \nVdash \Delta^2\neg\Delta^1 A$. But from (3) and the rules for Δ^1 and $\neg$: $s, N_1, M_2 \Vdash \neg\Delta^1 A$. Instantiating (S5$i$) for $i = 2$ therefore gives a conditional $\neg\Delta^1 A \rightarrow \Delta^2\neg\Delta^1 A$ whose antecedent is true at s, N_1, M_2 and whose consequent is not. Hence: $s, N_1, M_2 \nVdash \neg\Delta^1 A \rightarrow \Delta^2\neg\Delta^1 A$. So (S5$i$) is not supertrue in M_2 (for $i = 2$). The argument generalises to show that for no i is (S5i) supertrue on all i-models.

Williamson (1994, §5.6) defines an operator Δ^* as equivalent to an infinite conjunction:

$$\Delta^* A \text{ is true iff } \Delta^1 A \text{ is true, and } \Delta^2\Delta^1 A \text{ is true, and } \Delta^3\Delta^2\Delta^1 A \text{ is true, and} \ldots$$

[38] Note that (S4i) is not superfalse in M_2. From (3): $s, N_1, M_2 \nVdash \Delta^1 A$. So: $s, N_1, M_2 \Vdash \Delta^1 A \rightarrow \Delta^2\Delta^1 A$. So (S4$i$) is not superfalse in M_2 (for $i = 2$). Since the argument in the text shows that it isn't supertrue either: (S4i) falls down a supertruth-value gap in M_2 (for $i = 2$). In fact, (S4i) is not superfalse in any i-model. Similar remarks apply to (S5i).

The Δ^* cases represent the maximally clear cases: those without a hint of vagueness. Williamson claims that the following is valid:

(S4*) $$\Delta^* A \to \Delta^*\Delta^* A$$

He concludes that the maximally clear cases are sharply bounded. If so, then the Supertruth View places a logical limit on the extent of higher-order vagueness: borderline Δ^* cases are logically impossible. So UBC, and hence NSB, are false. But once one sharp boundary is accepted, what's wrong with more? Why is a sharp distinction between the Δ^* cases and the rest better than one between the cases and the non-cases? In other words: why not adopt an epistemic account of all vagueness, given that we have to do so for vagueness in Δ^*? Furthermore, since even a hint of unclarity suffices for untruth, the positive cases will be the Δ^* cases, and hence sharply distinguished from the rest. So there cannot really even be first-order borderline cases. The Supertruth View looks highly unstable.

Williamson's argument is not irresistible. Note first that the clause for Δ^*, unlike those for our Δ^i, employs a notion of truth without relativisation to any kind of model-structure. Thus Δ^* isn't well-defined in our framework. How can this be rectified? The most promising strategy combines 1-models, 2-models, 3-models, and i-models for every natural i into a single structure of the kind defined in the previous section. Truth in that kind of structure can then be used to supply truth-conditions for Δ^*:

$v, s, M_1, M_2, M_3 \ldots \Vdash \Delta^* A$ iff:

(i) $v, t, M_1, M_2, M_3 \ldots \Vdash A$, for all $t \in M_1$, and

(ii) $v, s, N_1, M_2, M_3 \ldots \Vdash A$, for all 1-models $N_1 \in M_1$, and

(iii) $v, s, M_1, N_2, M_3 \ldots \Vdash A$, for all 2-models $N_2 \in M_3$, and

$\vdots$

This validates (S4*). Does it show that Δ^* is precise, or that supervaluationist semantics imposes hidden sharp boundaries? A positive answer requires (a) that vagueness does not extend into transfinite orders, and (b) that this kind of model-structure can capture all the vagueness of a natural language with only finite orders of vagueness.

Set aside objections to (a): if each finite order of vagueness is captured by some iteration of our supervaluationist construction, then transfinite orders of vagueness should be captured by transfinite iterations. (Shapiro, 2006, ch.5.1 contains related discussion.) And even if not, the Supertruth theorist surely shouldn't *have* to appeal to something so recherché as transfinite orders of vagueness.

Assumption (b) is more dubious, and certainly not mandatory. We presented supervaluationism as a reasonable mathematical approximation to vague classification. The assumption that all the vagueness of a natural language can be captured without artefacts by a single mathematical structure is non-trivial. In fact, there are reasons independent of supervaluationism to doubt that it can be, and to which we now turn.

2.9.6 Sainsbury on vagueness and set-theoretic semantics

Mark Sainsbury argues thus:

> "Sets have sharp boundaries, or, if you prefer, are sharp objects: for any set, and any object, either the object quite definitely belongs to the set or else it quite definitely does not. Suppose there were a set of things of which "red" is true: it would be the set of red things. However, "red" is vague: there are objects of which it is neither the case that "red" is (definitely) true nor the case that "red" is definitely not true. Such an object would neither definitely belong to the set of red things nor definitely fail to belong to this set. But this is impossible, by the very nature of sets. Hence there is no set of red things." (Sainsbury, 1990, p.252)

For similar reasons, there can be no set of clearly red things, or clearly clearly red things, and so on. Granting Sainsbury's assumption about the sharpness of sets, it follows that no set-theoretic semantics can capture the vagueness of natural language. And if all mathematics can be captured within set-theory, then no mathematised semantics can capture vague classification without inaccuracy. Vagueness proper will be just what is missing from any such semantics; only a formal surrogate can remain.

This line of thought is attractive. But what exactly does it show? It does not show that mathematised semantics cannot provide insight into vague classification, only that it cannot *exhaust* vagueness. It remains open whether a given mathematical structure might closely resemble that of an appropriately circumscribed segment of vague classification. And that is all the Supertruth theorist need claim.

So long as our theoretical interest lies only in the structure of the F/non-F classification, we can rest content with considering individual sharpenings. Each will classify—*mis*classify, the Supertruth theorist will claim—some borderline cases one way or the other. But if we aren't interested in such close approximation—in distinguishing the clear from the borderline, and in assigning truth-values only to those (contentful) sentences that possess them—this need not undermine our employing a semantic theory based around individual sharpenings. Unless we are interested in vagueness, we need not require our semantic theory to respect, or even be capable of representing, vagueness-related truth-value gaps.

An interest in first-order borderline cases requires a different approach, based around 1-models and Δ^1. This improves on the classical semantic theory based around individual sharpenings by allowing expression of the claim that x is a borderline F, and thereby distinguishing the first-order borderline from the clear cases.

An interest in second-order vagueness requires a different approach again. The distinctions afforded by individual 1-models are too coarse for this. We need to consider 2-models and introduce Δ^2. This allows us to distinguish between cases, first-order borderline cases and second-order borderline cases. Think of this as an open-ended process. Moving up through new kinds of model provides better and better approximations, each capable of representing more aspects of vague classification than its predecessor. But we should not assume without argument that even the limit of this process will be perfectly accurate. And Sainsbury's argument provides positive reason to doubt that it will be, though without undermining the usefulness of supervaluationist semantics, provided its limitations are kept in mind.

On this approach, higher-order vagueness is analogous to indefinite extensibility:

> F is *indefinitely extensible* iff there is a function δ such that, for any collection x of Fs, $\delta(x)$ is an F that does not belong to x.

Although this lacks definite content without explication of the relevant (and somewhat murky) notions of function and collection, it will suffice for our purposes. Indefinitely extensible concepts are supposed to resist the formation of a collection that exhausts their instances: for any collection x, purported to be the collection of all Fs, $\delta(x)$ is an F that's not amongst x; hence x is not the collection of all Fs.

Likewise, a higher-order vague concept resists complete characterisation of its applicability and vagueness. Suppose we attempt to describe the applicability of vague F. A simple demarcation into the Fs and non-Fs is, at best, only borderline correct because it classifies some borderline cases one way or the other. So we introduce the notion of clarity. The sole purpose of this notion is to delimit the source of the original description's borderline status. But if F is second-order vague, the resulting description will also be only borderline correct because it counts some borderline clear Fs as clear Fs (for example). Unrestricted higher-order vagueness in F prevents complete description of the applicability and vagueness of F. Each attempted description can be no better than borderline correct. Exhausting the vagueness of F requires a clearly correct description of the ways in which all descriptions are only borderline correct. But since higher-order vagueness prevents any description from capturing the ways in which it is itself only borderline correct, it's impossible to exhaustively describe the vagueness of F. (This meshes nicely with the view on which the ineradicable fuzziness of NSB implies, but is not analysed by, the higher-order borderline cases of UBC.)

If this line of thought is correct, then we may have an explanation of why no set-theoretic structure is vague: the nature of a set is exhausted by a list of its members, but no list of its instances can exhaust the applicability of a vague predicate, regardless of how fine-grained the distinctions we make amongst items on the list.

The present approach differs from a prominent alternative: insist that an adequate specification of the sharpenings requires a vague metalanguage (Keefe, 2000, ch.8 §1). Vagueness in the specification of sharpenings can induce vagueness in whether a sentence is true at them all, and hence vagueness in the truth-status of claims about clarity. Our approach, by contrast, makes no appeal to a vague metalanguage. Instead, we accept an in-principle limit on how closely our model can approximate vague classification. Our Supertruth theorist's metalanguage is the standard (and precise) language of classical mathematics. This affords a significant advantage over the alternative: we do not need to know what forms of reasoning are valid in a vague language before our investigation begins. Were our metalanguage vague, we would need to know the effect of vagueness on validity before we could derive any results about the model. But studying validity is just what we want the model for. So we use a standard mathematical metalanguage to approximate vagueness as best we can.

2.9.7 Objection: the fragmentation of vagueness

The present treatment distinguishes Δ^1 and Δ^2 by their semantic axioms and the kinds of structure those axioms presuppose. This section considers the objection that this misrepresents the unitary nature of clarity: our proposal on behalf of the Supertruth theorist breaks clarity into a cluster of formally similar distinct concepts, thereby misrepresenting vagueness as a non-uniform phenomenon.

This is not compelling. We distinguished different orders of borderline case and the kinds of structure needed to represent those orders. But within any model of any orders of vagueness, vagueness is represented by the structure as a whole, not any particular component of it. We could even, if we wished, define a single Δ operator capable of capturing all the orders of vagueness represented by a single model-structure.[39]

Furthermore, our approach is formally equivalent to a more common one, against which the objection is without force. This alternative introduces an accessibility relation on a single space of sharpenings and characterises a single clarity operator via truth at all accessible sharpenings, instead of at all sharpenings. Both approaches impose a hierarchical structure on a space of sharpenings and treat higher-orders of vagueness in terms of higher levels in the hierarchy. On the alter-

[39] Whether truth-value across sharpenings, 1-models, 2-models and so on was relevant to the truth of a sentence featuring this operator would depend on how many other occurrences of it occurred within its scope.

native however, there is no temptation to treat clarity as non-uniform. Given this formal equivalence, that temptation should not arise on our approach either.

Given this equivalence and the alternative's greater elegance, why bother with our approach at all? The answer is that it has a philosophical benefit that the alternative lacks. On our approach, consideration of higher and higher orders of vagueness requires that we consider different kinds of structure and define our semantic axioms anew for each one. Although inconvenient, this reminder of our theory's representational limit serves as a warning against assuming the meaningfulness of operators like Δ^*, which assume a complete hierarchy of orders within a single model. It also warns against assuming the possibility of capturing all the vagueness of natural language within a single mathematical structure; it warns against assuming that clarity in our representation is always indicative of clarity in the system it represents.

2.9.8 More hidden sharp boundaries?

Shapiro (2006, p.128) presents an argument similar to Williamson's Δ^* argument, but that makes no explicit assumptions about the semantics of vagueness and is therefore immune to our response to Williamson's argument.

Consider the absolutely clear *F*'s: the *F*'s about whose *F*-ness there isn't even the slightest hint of unclarity. Absolute clarity is the informal analogue of Δ^*. Suppose that *a* is borderline absolutely clearly *F*. Then there is a hint of unclarity about *a*'s *F*-ness. So *a* is not absolutely *F*. But if *a* is not absolutely *F*, then it's not borderline whether *a* is absolutely *F*. Since *a* was arbitrary and this rests on no assumptions: borderline absolutely clear *F*'s are impossible; the absolute *F*'s *must* be sharply bounded.

This is not uncontroversially valid. Let's use ABS for an absolute clarity operator. The argument began by supposing that *a* is borderline absolutely clearly *F*:

$$\neg\Delta\,\textsc{abs}\,Fa \wedge \neg\Delta\neg\,\textsc{abs}\,Fa \tag{4}$$

The task is to show that borderline absolutely clear *F*s are impossible, and hence that (4) is false. Shapiro's argument begins by inferring from (4) that *a* is not absolutely clearly F on the grounds that there is a hint of unclarity about its F-ness:

(5) $$\neg \textsc{abs}\, Fa$$

The conjunction of (4) and (5) is not a contradiction. So we can't yet conclude that (4) is false, or that a is not a borderline absolutely clear F:

$$\neg(\neg\Delta \textsc{abs}\, Fa \wedge \neg\Delta\neg \textsc{abs}\, Fa)$$

How might we get a contradiction? Consider a version of the S5 axiom for ABS:

(S5ABS) $$\neg \textsc{abs}\, A \rightarrow \textsc{abs} \neg \textsc{abs}\, A$$

From (5), this yields:

$$\textsc{abs} \neg \textsc{abs}\, Fa$$

Then because ABS A implies ΔA:

$$\Delta\neg \textsc{abs}\, Fa$$

This contradicts the second conjunct of (4). But we've already seen that principles like (S5ABS) shouldn't be unrestrictedly valid, if we're going to allow for higher-order vagueness.

An alternative strategy appeals to: $A \models_{\text{global}} \Delta A$. From (5), this yields

(6) $$\Delta\neg \textsc{abs}\, Fa$$

Which again contradicts the second conjunct of (4). But $A \models_{\text{global}} \Delta A$ requires only that ΔA is (super)true in any model where A is (super)true. This is silent about the (super)truth-status of ΔA when A is borderline and hence untrue. It's therefore silent about the (super)truth-status of (6) under the supposition that (5) is borderline. Yet that's just what (4) implies (since $\neg A$ is borderline whenever A is borderline). So on the Supertruth View, it follows only that if it's (super)true, and hence clearly true, that a is not absolutely clearly F, then its's not borderline whether a is absolutely clearly F. This is obviously unhelpful when trying to reduce the supposition that a is borderline absolutely clearly F to absurdity.

This response is problematic. ABS Fa is false if (5) is true. But (4) says that ABS Fa is borderline, and hence valueless. Since nothing can be both false and valueless,

(4) and (5) cannot both be true. So (4) cannot be true, if it implies (5). So borderline absolutely clear cases are impossible after all.

The Supertruth theorist must therefore reject the initial step from (4) to (5). Although (4) implies that ABS Fa isn't true, it's negation need not be true: ABS Fa can fail to be true by being borderline, just as (4) says, without thereby being false, as (5) says.

Given the Supertruth View's connection between borderline status and truth-value gaps, the argument from ⌜Borderline-A⌝ to ⌜Not-A⌝ shouldn't be valid. Shapiro's argument therefore does not show that borderline absolutely clear cases are impossible. It does show that all borderline absolutely clear cases fail to be absolutely clear cases. But that's compatible with their failing to be non-cases of absolutely clear cases too, given the identification of borderline status with truth-value gaps. If an object can satisfy neither A nor $\neg A$, as the Supertruth theorist claims, there seems no reason why it couldn't satisfy neither ABS A nor $\neg$ ABS A. Yet the Supertruth theorist must provide an account of clarity that explains the compatibility of failure to be a case with being a borderline case, as opposed to a non-case. Since the borderline cases fall down a truth-value gap, this requires an account of truth that distinguishes untruth from falsity. Without that account, the present line of resistance to Shapiro's argument looks more like wishful thinking than a principled response. Since that explanation is just what the Supertruth theorist has yet to provide—recall the discussion in §2.7.1—their position is tenuous. The difficulty explicating their non-Bivalent conception of truth undermines the Supertruth theorist's ability to respond to Shapiro's argument for hidden sharp boundaries.

A better supervaluationist strategy would be to find an alternative response. There seem to be three options. (i) Deny that the concept of absolute clarity is coherent. I can see no good reason to grant this. (ii) Accept this limit on higher-order vagueness, but resist positing sharp boundaries elsewhere. I don't know how to do so in a principled manner.[40] (iii) Deny that there are any absolutely clear cases. The

[40] One strategy might claim that absolute clarity is a theoretical concept and so deny that the intuitive reasons to acknowledge borderline cases apply to it: the absolutely clear F's are not sorites-susceptible because there's no extra-theoretical motivation to grant a sorites premiss for 'absolutely

next section argues that the Sharpening View can motivate this denial. I know of no other way to do so. But without such motivation, response (iii) looks objectionably *ad-hoc*. Hence it is doubtful whether the Supertruth View can accommodate unrestricted higher-order vagueness.

2.9.9 Higher-order vagueness and the Sharpening View

The Sharpening View fares better with higher-order vagueness. The view is motivated by a picture of Reality as a gradual place; vagueness arises when we impose non-gradual classifications upon it. This gradualness and the relative coarseness of the meaning-determining facts combine to ensure that no one classificatory boundary is privileged over all others, despite many being ruled out. Each remaining boundary provides an intended interpretation of the expression in question.

We can apply this to metasemantic vocabulary. The result is a well-motivated denial that there are any (typical) absolutely clear or Δ^* cases. Hence neither Williamson nor Shapiro's argument to show that such cases are sharply bounded shows that actual vague classification is sharply bounded.

Before we begin, it's worth responding to the following objection: surely there are absolutely clear cases; isn't scarlet as clearly a shade of red as anything could be? The approach below attempts to offset the strangeness of the claim that scarlet isn't absolutely clearly a shade of red by allowing that it's as clearly a shade of red as anything could be, given our limitations and the way we use language. Extremely clear cases are commonplace and susceptible to borderline cases, but absolutely clear cases are not. The Sharpening theorist's response is to accuse the objector of confusing absolute clarity with very high levels of clarity.

2.9.9.1 Metasemantic gradualness

We want metasemantic vocabulary to be vague. So we need to show how the facts we describe using that vocabulary can be gradual, just like the facts described by typical vague vocabulary.

clear *F*' or 'not absolutely clear *F*'. But although absolute clarity is a theoretical concept, it doesn't seem so far divorced from ordinary clarity that we shouldn't find attractive either Sorites premisses for it, or the claim that it permits borderline cases. Thanks to Will Bynoe for suggesting this strategy.

The metasemantic facts impose an ordering on interpretations according to how well they fit a community's linguistic behaviour (§2.4.2). For example, an interpretation that places the tall/non-tall distinction at 5'11" fits our use of 'tall' better than one that places it at 5'10", but (perhaps) less well than one that places it at 6'. Small differences between heights bring small differences in how well interpretations that locate the tall/non-tall boundary at those heights fit our use of 'tall'. Gradualness in the height-facts underlying our use of 'tall' thus induces gradualness in the metasemantic facts underlying our use of 'intended interpretation' (and vagueness in object-language reflections thereof, like 'said that').

This gives an ordering on interpretations according to how well they fit a community's use of language. Intended interpretations are greatest elements in this ordering. But does the class of such greatest elements provide a significantly better interpretation of 'intended interpretation' than any other? Or, like the distinction between interpretations of 'tall' that place the tall/non-tall boundary at 6' and those that place it at 6'0.00001", is this a distinction without a difference? In the former case, metasemantic vocabulary will be non-vague. In the latter case, metasemantic vagueness seems likely: no one point in the gradual fit-transition described by non-gradual metasemantic vocabulary is significantly better than all others. Answering this question requires an account of the nature of the facts underlying discourse about intended interpretations.

Distinguish two broad approaches to metasemantics. One sees metasemantic facts as a *sui generis* kind of fact, though systematically connected to other kinds of fact, such as those about language-use. On this view, the intended/unintended distinction is naturally taken to mark a significant difference.[41] This isn't threatened by the fact that the intended/unintended distinction isn't revealed in a description of the fit-ordering alone; for that distinction isn't supposed to be reducible to facts about fit, despite coinciding with the greatest/not-greatest distinction in the fit-ordering.

The alternative approach sees metasemantic discourse as codifying other facts

[41] A conception of semantic facts as *sui generis* only makes it natural, not mandatory, to regard the intended/unintended distinction as significant. For that distinction could be a *sui generis* gradual distinction. It's hard to find a motivation for this view.

about, say, the relevance of various propositional contents to linguistic communication within a community. This kind of approach undermines the significance of the intended/unintended distinction by denigrating the greatest/not-greatest distinction in the fit-ordering; although that distinction exists, it marks no significant difference. What really matters is not which interpretations are greatest in the fit-ordering, but the ordering itself.

On this second approach, the class of greatest interpretations in the fit-ordering need not provide a significantly better extension for 'intended interpretation' than a more inclusive class containing some marginally less well fitting interpretations. Metasemantic theorising mandates drawing a distinction somewhere in the fit-ordering, though no one candidate is significantly better than all others. Vagueness can then infect metasemantic concepts just as it infects any others. In this case, it can be vague which sentences are clearly true.

This suggests a philosophical interpretation of the modified supervaluationist formalism described in §2.9.4. Sharpenings represent interpretations of the non-metasemantic vocabulary. 1-models represent classes of such interpretations: interpretations of 'intended interpretation'. 2-models thus represent classes of interpretations of 'intended interpretation'. By parity of reasoning, 2-models represent interpretations of 'intended interpretation of 'intended interpretation' '; they represent states of the metasemantic facts on which there are many intended interpretations of 'intended interpretation of the non-metasemantic vocabulary'. Likewise *mutatis mutandis* for 3-models and above.

Think of 'intended interpretation' as marking a threshold in the fit-ordering: interpretations that fit better than the threshold count as intended. Objects that satisfy F under each interpretation that exceeds this threshold are clearly F (under that interpretation of 'intended interpretation'). Thresholds in the fit-ordering correspond to 1-models.

Now, if the metasemantic facts are gradual—i.e. if nearby interpretations in the fit-ordering fit the meaning-determining facts almost as well as each other—then many nearby thresholds will have equal claim to be the threshold marked by 'intended interpretation': none marks any significant difference in how well interpretations that exceed that threshold fit the meaning-determining facts, or their relevance to linguistic communication within the community in question. Objects that satisfy F under each interpretation that meets each of these thresholds are clearly clearly F. Classes of thresholds in the fit-ordering correspond to 2-models. The class of interpretations that meet any of these thresholds corresponds to the class of sharpenings that belong to any of the 1-models within a 2-model.

Provided the metasemantic facts are sufficiently gradual, this process should iterate: many classes of classes thresholds in the fit-ordering will have equally good claim to be the intended interpretation of 'intended interpretation of 'intended interpretation' ', and so on. The key point is that iterating 'clearly' to force consideration of new kinds of model slightly reduces the threshold in the fit-ordering that determines the class of interpretations such that an object must satisfy F under each member of that class in order to count as clearly…clearly F. The clearly clear Fs satisfy F under a more inclusive class of interpretations that do the (mere) clear Fs; and the clearly clearly clear Fs satisfy F under a more inclusive class of interpretations still. In each case, the more inclusive class contains those interpretations that fit only slightly less well than the interpretations in the less inclusive class. Limits on metasemantic gradualness will limit higher-order vagueness by limiting how inclusive a class of interpretations can be obtained by successive iterations of 'clearly'.

2.9.9.2 Absolute clarity

This kind of view provides reason to deny the existence of Δ^* or absolutely clear Fs. Pre-fixing a sentence with another occurrence of Δ requires truth under a more inclusive classes of interpretations. The interpretations in successively more inclusive classes fit our use of language only slightly less well than do those in the immediately preceding less inclusive class. If any interpretation is connected to any other interpretation by a series of only slightly less well fitting interpretations, then the absolutely clear Fs will be the objects that satisfy F under every interpretation. Since nothing does so, there will be no absolutely clear of Δ^* Fs. So even if Δ^* and absolute clarity are precise, our actual vague classification won't be.

Why think that the metasemantic facts are like this? Mightn't some interpretations just be utterly irrelevant to our communication with the vague predicate F? Maybe so. But that's not what's at issue. The question is whether the interpretations that count x as an F are connected to those that don't do so by a series of successively less well-fitting interpretations, not whether such interpretations are utterly irrelevant. If there's no such series, then x will be absolutely clearly F. The class of such objects will be the class of absolutely clear Fs, and there will be no borderline cases to this class.

This kind of limit on higher-order vagueness differs from those argued for by Williamson and Shapiro. Their arguments would make it logically impossible for a language to lack sharp boundaries entirely. The present kind of restriction on higher-order vagueness results from contingent features of the ordering of interpretations by fit.

Does our use of ordinary vague predicates give rise to such disconnected series' of interpretations? (And if it does, are the disconnected interpretations the only ones that count certain objects one way or the other?) That's what's required for the present kind of limit on higher-order vagueness. It's hard to believe that our use of language is like this. Small differences in the respects to which our use of F is sensitive correlate with small differences in whether typical speakers would, by and large, judge the objects in question to be F. And it is those judgements that are primarily responsible for how well an interpretation fits our use of F. I can see only one way of introducing a predicate that would generate discontinuities in the fit-ordering of the kind necessary to restrict higher-order vagueness.

The method I have in mind introduces predicates by ostending a determinate range of paradigm cases: *this*, *that*, *the other* and anything sufficiently similar to them are all and only the Fs. Vagueness in 'similar enough' induces vagueness in F. But any interpretation that places *this*, *that* and *the other* outside the extension of F is utterly irrelevant to this use of F, and significantly less relevant than all other interpretations. So *this*, *that* and *the other* are absolutely clearly F because no Sorites series of gradually less well-fitting interpretations connects one that places the paradigms outside the extension of F to one that places them inside it. Yet for any other object, there may well be such a series. Hence *this*, *that* and *the other* will be all and only the absolutely clear Fs. Here we have a sharp boundary resulting from our use of F.

This kind of case is atypical. Ordinary concepts are not introduced by ostension of a determinate range of paradigms. As soon as a condition expressed by an ordinary predicate is used in place of ostension when determining the paradigms, an appropriate sorites series of interpretations will result, and so there will be no absolutely clear cases. Our Lewisian Sharpening View thus combines with a reductive approach to metasemantics to generate a response to Williamson and Shapiro's arguments for hidden sharp boundaries in our actual vague classification.

2.9.10 Higher-order vagueness: concluding remarks

Does the Sharpening View or the Supertruth View provide the better approach to higher-order vagueness? Well, the Supertruth View faces three problems.

Firstly, the Supertruth View's account of higher-order vagueness undermines its analysis of clarity in terms of truth and falsehood. Since all orders of borderline case fall down a truth-value gap, there are distinctions marked by Δ that cannot be explained in terms of truth, falsity and gaps.

Secondly, the response to Shapiro's argument requires supplementation with an account of a non-Bivalent notion of truth (that supports LEM). §2.7.1 argued that it is doubtful whether this is possible.

Thirdly, the response to Williamson's argument for sharp boundaries requires accepting a limit on how accurate the semantic theory can be. Although we shouldn't assume that a perfectly accurate mathematised semantic theory will be possible, a view that purports to offer one is *ceteris paribus* preferable to one that does not.

None of these difficulties afflicts the Supertruth View. So that view is preferable.

2.10 Conclusion

We began with a formal setting and a range of philosophical interpretations that might be imposed upon it. All bar two were ruled out in §§2.3–2.4. We then investigated a range of difficulties for these two remainders. In each case, the Sharpening View was seen to be less problematic than the Supertruth View. We also saw that adequate responses to most problems with the Supertruth View require an account of its non-Bivalent conception of truth. It is not clear what this account might look like.

The problems with explaining the Supertruth theorist's conception of truth all stem from the identification of clear truth with truth. This suggests that what the Supertruth View really lacks is not so much an account of truth, but an account of vagueness: the challenge of providing an account of clarity has simply been transformed into the challenge of providing an account of truth. Hence the remainder of this thesis will focus on the Sharpening View. We will, however, highlight those points where the Supertruth View makes a difference (the next chapter's discussion of vague reference will contain quite a few).

One final question: to what extent is the Sharpening View a version of supervaluationism? Keefe does not think it is. She calls it (or something very much like it) the "pragmatic theory of vagueness" and opposes it to her own supervaluationism (which is itself a version of our Supertruth View) (Keefe, 2000, ch.6). The dispute is terminological. The views share (i) a formal structure, (ii) an analysis of clear truth as supertruth, and (iii) a non-privileging of any one sharpening over any other in the apparatus of truth-evaluation. The differences concern only the metaphysics of models and sharpenings, and whether the primary notion of semantic evaluation is supertruth or s-truth. Furthermore, our Sharpening View, unlike the Supertruth View does justice to the idea of vagueness as "semantic indecision" or under-determination of content often associated with supervaluationism (§2.3.3.3), and which Keefe herself endorses. It is therefore not misleading to describe the Sharpening View as a form of supervaluationism.

Chapter 3

Vagueness in Reference

This chapter and the next examine different applications of the Sharpening View developed in the preceding chapter to the Problem of the Many. This chapter examines the idea that an ordinary object's boundaries are vague insofar as it's vague which individual's boundaries are at issue: Tibbles's boundaries are vague because it's vague which object 'Tibbles' refers to. On this approach, Unger's puzzle becomes a source of referential unclarity, and hence also of unclear boundaries. This kind of view will ultimately be rejected.

This chapter also serves a second purpose: to defend and elaborate our Sharpening View in response to several objections. These objections don't concern the Problem of the Many, so much as supervaluationist accounts of vague reference. Since there may be sources of referential vagueness other than the Problem of the Many, a full defence of the Sharpening View must address these objections.

Lewis endorses this reduction of vague boundaries to referential vagueness, and describes two ways of applying it to the Problem of the Many. §3.1 begins by presenting both and rejecting one. §3.2 then examines four objections to supervaluationist accounts of referential vagueness. These concern: indirect reports of vague speech; violation of a plausible constraint on reference; *de re* thought; conflict with Direct Reference theory. Each objection will be found wanting. Three more serious difficulties for the Lewisian account of vague boundaries itself are presented in §3.3. Doubts will be raised about whether this approach: provides a genuine solution to the Problem of the Many, or merely makes it difficult to express it; can accommodate vagueness in the boundaries of self-referrers; is separable from objectionable components of Lewis's metaphysical system. §3.4 concludes.

3.1 Two solutions

This chapter examines the following Proposal in the context of the Sharpening View of vagueness that we developed in chapter 2:

> Vagueness in the boundaries of ordinary objects results from vagueness about which individual's boundaries are in question,

Lewis (1993a) develops this Proposal in two ways. This section presents both and rejects one.

3.1.1 Two options

This section introduces the Proposal in a little more detail.

Let h be one of Tibbles's borderline hairs:

$$\neg\Delta h \text{ is part of Tibbles } \wedge \neg\Delta\neg h \text{ is part of Tibbles.}$$

It's not plausible that vagueness in 'h' is responsible for this. So given the Sharpening View's account of vagueness of as multiplicity of interpretation, there are two options:[1]

(i) There are many intended interpretations of 'Tibbles'.

(ii) There are many intended interpretations of 'is part of'.

This chapter examines (i). The next chapter examines (ii).

Option (i) traces vagueness in Tibbles's boundaries to vagueness about which object 'Tibbles' refers to. This is Lewis's view:

> "The only intelligible account of vagueness locates it in our thought and language. The reason it's vague where the outback begins is not that there's this thing, the outback, with imprecise borders; rather, there are

[1] Although these options could be combined, it's (a) obscure why we would want to, and (b) clearer to discuss them separately.

> many things, with different borders, and nobody has been fool enough to try to enforce a choice of one of them as the official referent of the word 'outback'. Vagueness is semantic indecision." (Lewis, 1986b, p.212)

Note that the argument from:

> Vagueness is a feature of our thought and language.

and:

> Vagueness is semantic indecision.

to:

> Vague boundaries are the product of referential vagueness

is invalid without further premisses to rule out option (ii) above. One can consistently endorse the first two theses despite rejecting the third. The next chapter defends a view that does just that.

We focus on mereological vagueness, rather than vague boundaries or locations. We lose no generality because (i) analogous remarks apply to these other kinds of vagueness, and (ii) mereological vagueness implies vague boundaries and locations via:

> x is located in region r iff some part of x is located in region r.

So suppose h is Tibbles's only borderline part. Let T^+ be the fusion of h with the rest of Tibbles. Let T^- be T^+ excluding h. Which of T^+ and T^- is Tibbles?[2] And which is a cat? Both are equally good candidates. The Lewisian Proposal is that our use of 'Tibbles' doesn't distinguish between them: each is an intended referent of that name. Formally, this amounts to allowing supervaluationist models M that violate the following constraint we imposed in §2.1:

> For any sharpenings $s, t \in M$ and singular term $\alpha : [\![\alpha]\!]_s = [\![\alpha]\!]_t$.

Now, four desiderata:

[2] We needn't assume that either T^+ or T^- is Tibbles. We could instead ask which constitutes Tibbles, or take them as collections of microscopic particles and ask which composes Tibbles. The formulation in the text simply aids presentation. See §1.1.2.2 for discussion.

(D1) This should be borderline: 'h is part of Tibbles'.

(D2) This should be clearly true: 'Tibbles is a cat'.

(D3) This should be clearly true: 'There is exactly one cat on the mat'.

(D4) This should be borderline: 'h is part of the cat on the mat'.

Since h is part of T^{+} but not of T^{-}, cross-sharpening variation about the referent of 'Tibbles' secures (D1). Both approaches below agree on this. They diverge over (D2)–(D4) and the treatment of 'cat'. In the following, our quantifiers will often be tacitly restricted to objects on the mat, thereby allowing us to drop the qualification 'on the mat': there is exactly one cat, and it's borderline whether h is part of it.

Is there an argument for this Lewisian Proposal? One strategy compares it to its rivals, arguing that, on balance, the Proposal is preferable. However, a more satisfying approach would begin with an account of ordinary objects, or of material reality more generally, that implies (or at least suggests) the Proposal; we would like something more than an *ad-hoc* collection of theses unified only by their ability to solve certain problems. I know of two closely related such accounts.

The first is Quine's: a material objects is just "the material content of any portion of space-time, however scattered and discontinuous" (Quine, 1976).[3] Precision in the boundaries of regions of space-time translates into precision in the boundaries of material objects. Vagueness in claims about the boundaries of those objects must therefore come from vagueness as to their subjects.

The second account appeals to Lewisian views about the logicality and "ontological innocence" of classical extensional mereology, and his use of that mereology in the foundations of set-theory (Lewis, 1991, 1993b). Since logical and (pure) mathematical vocabulary cannot, it seems, be vague, vagueness in mereological predications must result from vagueness as to their subjects.

Neither Quine's nor Lewis's view carries intuitive force. Furthermore, the close connections between the Proposal and other elements of Lewis's metaphysical system are already beginning to emerge. Our later arguments against the Proposal (§3.3) therefore also count against these elements of Lewis's system.[4]

[3] A related view reduces objects to the regions of spacetime they occupy.

[4] §3.3.3 argues from the Proposal to counterpart-theory and perdurance. We then argue against

3.1.2 Unger and Lewis

What is the relationship between Unger's puzzle of too many candidates and Lewis's puzzle of vagueness and borderline candidates, given the Proposal? When there are many best (and good enough) near-coincident cat-candidates on the mat, our use of 'Tibbles' won't distinguish amongst them. Each will be an intended referent of 'Tibbles'. The result is referential unclarity in 'Tibbles' and unclarity about Tibbles's boundaries. Unger's puzzle thus becomes a source referential and mereological unclarity.

It doesn't follow that Unger's puzzle is a source of referential vagueness because it doesn't follow that 'is part of Tibbles' is Sorites-susceptible or admits of higher-order borderline cases. That seems to require the gradualness of boundary-transition that motivates Lewis's puzzle (§1.2). But the end result is the same: our use of 'Tibbles' isn't fine-grained enough to distinguish one from amongst a range of candidates on the mat. Since each candidate fits our use equally well (and well enough), each is an intended referent for 'Tibbles' and contributes to unclarity in mereological sentences featuring that name.

Unger argues that his and Lewis's puzzles are distinct because his arises even under the supposition that Tibbles's boundaries are entirely precise (§1.3). On the present approach, his reasoning is flawed. If Tibbles's boundaries are precise, then there's a unique intended referent for 'Tibbles', and Unger's puzzle doesn't arise. For if 'Tibbles' has a unique intended referent, then there's a unique most cat-like object on the mat. Were there several such objects, our use of 'Tibbles' wouldn't distinguish between them; so 'Tibbles' wouldn't have a unique intended referent; so Tibbles's boundaries wouldn't be precise. It is therefore safe to focus on Lewis's puzzle in the remainder.

3.1.3 One Cat

This section presents Lewis's first account of mereological vagueness and the Problem of the Many. It combines two theses. The first concerns the extension of 'cat':

---

the Proposal by arguing against these views. Arguments against counterpart-theory and perdurance thus translate into arguments against Quine's and Lewis's views about objects and mereology.

> Exactly one Tibbles-candidates satisfies 'cat' at each sharpening; different candidates at different sharpenings (and each candidate at some sharpening).

Desideratum (D3) holds because this makes it supertrue that there's exactly one cat on the mat. Note however, that there's nothing of which it's supertrue that it is a cat. Desideratum (D4) holds because 'h is part of the cat' is s-true iff the s-satisfier of 'cat' includes h, and not all candidates do so.

The second thesis posits a penumbral connection between 'cat' and 'Tibbles':

> For each sharpening s : $[\![\text{'Tibbles'}]\!]_s \in [\![\text{'cat'}]\!]_s$.

Desideratum (D2) holds because this makes it supertrue that Tibbles is a cat. This constraint transfers referential vagueness in 'Tibbles', and hence mereological vagueness in Tibbles, to predicative vagueness in 'cat'. This is as it should be: 'Tibbles' was introduced as a name for an individual cat.

We will call this the One Cat (OC) solution. On this approach, vague boundaries reflect referential vagueness in names for ordinary objects. This referential vagueness induces corresponding predicative vagueness in our ordinary sortal concepts.

3.1.4 Many Cats

This section presents Lewis's second account of mereological vagueness and the Problem of the Many. The underlying idea is that every sufficiently cat-like object counts as a cat. The Tibbles-candidates are all sufficiently cat-like. So they are all cats: each satisfies 'cat' under each sharpening. Hence, for each sharpening s, the s-referent of 'Tibbles' belongs to the extension of 'cat'. So it's supertrue, and hence clearly true, that Tibbles is a cat. So desideratum (D2) is satisfied.

Let 'Cx' formalise 'x is a cat'. Then the following is supertrue when restricted to objects on the mat:

(1) $$\exists x \exists y (Cx \wedge Cy \wedge x \neq y)$$

But recall desideratum (D3): it should be clearly true that there's exactly one cat. So Lewis must deny that (1) expresses the truth-condition of the English 'there are (at least) two cats'. An alternative truth-condition is required.

Let '$\textsc{nco}(x, y)$' formalise 'x and y nearly materially coincide with one another'. Lewis (1993a, p.178) suggests the following truth-condition for English numerical claims 'there are(at least) n F's':

$$\exists x_1 \ldots \exists x_n (Fx_1 \wedge \ldots \wedge Fx_n \wedge \neg\textsc{nco}(x_1, x_2) \wedge \neg\textsc{nco}(x_1, x_3) \wedge \ldots \wedge \neg\textsc{nco}(x_{n-1}, x_n))$$

English individuative vocabulary thus gets interpreted using near-coincidence rather than numerical identity. Since T^+ and T^- are the only Tibbles-candidates, this is supertrue:

$$\exists x \forall y (Cx \wedge (Cy \rightarrow \textsc{nco}(x, y))$$

So 'there is exactly one cat' is supertrue, despite both T^+ and T^- satisfying 'cat'. Hence (D3) is satisfied. Note also that, unlike the OC approach, it's supertrue of each cat-candidate that it's a cat.

Desideratum (D4) is trickier. This should be borderline:

h is part of the cat.

Let '$x \leq y$' formalise 'x is a (proper or improper) part of y'. Then we have two Russellian truth-conditions:

$$\exists x \forall y (Cx \wedge [Cy \rightarrow y = x] \wedge h \leq x)$$
$$\exists x \forall y (Cx \wedge [Cy \rightarrow (y = x \wedge h \leq y)])$$

Although equivalent and supertruth-valueless on the OC approach, the present approach makes them non-equivalent:

$$\exists x \forall y (Cx \wedge [Cy \rightarrow \textsc{nco}(y, x)] \wedge h \leq x)$$
$$\exists x \forall y (Cx \wedge [Cy \rightarrow (\textsc{nco}(y, x) \wedge h \leq y)])$$

The first is supertrue because every candidate/cat nearly coincides with a cat of which h is part, namely T^+. The latter is superfalse because h is not part of every candidate/cat; specifically, it is not part of T^-. Since it should be borderline whether h is part of the cat, neither truth-condition is correct.

Lewis suggests two ways around this problem. The first treats descriptions as singular terms, rather than disguised quantifier phrases: vagueness in 'h is part of the cat' is just like vagueness in 'h is part of Tibbles'. Lewis's second suggestion is that both formalisations express intended interpretations of 'h is part of the cat': our use of definite descriptions doesn't distinguish between these truth-conditions. This brings cross-sharpening variation in truth-value, thereby making it borderline whether h is part of the cat. We should however, be sceptical of this second solution if a quantificational treatment of descriptions is attractive. The reason is that it prevents a unitary analysis of vague boundaries: vagueness about Tibbles's boundaries results from referential vagueness in 'Tibbles', while vagueness about the cat's boundaries results from vagueness in the truth-conditions of descriptions.

Like the OC approach, this Many Cat (MC) solution sees vague boundaries as a reflection of referential vagueness (modulo the worries at the end of the preceding paragraph). Unlike the OC approach however, this referential vagueness doesn't bring predicative vagueness in ordinary sortals, but a non-standard interpretation of individuative vocabulary.

So, Lewis offers two ways to maintain that Tibbles is the only cat on the mat, despite his vague boundaries. Both postulate referential vagueness in 'Tibbles'. And both employ supervaluations to make it borderline whether h is part of Tibbles. They differ in two ways. Firstly, over the interpretation of 'cat': does it s-apply only to the s-referent of 'Tibbles', or to every sufficiently cat-like object on the mat? Secondly, over the interpretation of English individuative vocabulary: does it express identity and distinctness or near-coincidence and (extensive) disjointness?

Which approach is preferable? Lewis endorses both, arguing that different contexts require different solutions. The next section argues that he is wrong, and the MC approach should be rejected.[5]

[5] Williams (2006) offers a positive argument for the MC approach. His argument relies on attributing the following conjecture to supervaluationists: ordinary speakers reason as if (clearly) true existentials require (clearly) true instantiations. He claims that this is required by the best supervaluationist response to the Sorites. But our Sharpening-theoretic response to the Sorites in §2.5.2 rejected that thesis in favour of an alternative. Williams's argument for the MC approach is therefore without force in the present context.

3.1.5 The Problem of the Two

We've got two approaches to the Problem of the Many in place. Isn't this one too many? Lewis thinks not. He claims that different contexts require different solutions, and hence that it's context-sensitive which solution applies. This section argues that he is wrong.

Lewis (1993a, p.180) claims that the MC approach is required when we discuss vagueness because attending to the equally cat-like natures of the candidates places them all in the extension of 'cat'. This can't be right; for Lewis also discusses a case that the MC approach cannot accommodate, regardless of whether we're discussing vagueness. It follows that even if the choice between OC and MC solutions is context-sensitive, it's not sensitive to the difference between contexts in which we're discussing vagueness and more typical contexts (outside the philosophy seminar): the MC approach cannot apply in every member of either class of contexts. No such problem afflicts the OC approach. So unless it is context-sensitive which solution applies, the MC approach fails and only the OC approach is defensible. The defender of the MC approach therefore requires an alternative account of when it applies. None is forthcoming.

If this right, then Lewis's postulated context-sensitivity brings two problems. Firstly, it multiplies senses of common nouns and individuative vocabulary without necessity. Secondly, it undermines our semantic theory's systematicity by positing context-sensitivity without an account of which features of context the sensitivity is to, or why. Since the MC solution is defensible only by appeal to context-sensitivity, it ought therefore to be rejected.

Here's how Lewis presents the problematic case:

> "Fred's house taken as including the garage, and taken as not including the garage, have equal claim to be his house. The claim had better be good enough, else he has no house. So Fred has two houses. No!" (Lewis, 1993a, pp.180–1)

Since the candidates don't nearly coincide, the MC approach makes 'Fred has two houses' supertrue. But whether we're admiring Fred's garden or discussing the semantics of vagueness, that sentence ought to be false. So the MC approach doesn't apply to all typical contexts, and it doesn't apply whenever we're discussing vagueness.[6] So when does it apply? Without a well-motivated answer to this question we ought to reject the MC approach. So let us do so.

3.2 Four problems with vague reference

This section rebuts four objections to supervaluationist accounts of vague reference, and hence also (indirectly) to the Lewisian Proposal about the Problem of the Many. The objections concern: indirect speech reports in §3.2.1; a plausible constraint on reference in §3.2.2; *de re* thought in §3.2.3; Direct Reference in §3.2.4. §3.3 turns to three more serious worries for the Lewisian approach to the Problem of the Many itself.

3.2.1 Schiffer on speech reports

Stephen Schiffer (1998, §1; 2000, pp.321–6) argues that supervaluationism makes indirect reports of vague speech false. §3.2.1.1 presents three problem cases. §3.2.1.2 presents a simple semantics for indirect reports and a diagnosis of the problem. §3.2.1.3 responds to Schiffer by implementing this semantics within our Sharpening View. §3.2.1.4 addresses another difficulty Schiffer raises for this kind of approach.

3.2.1.1 Three problem cases

This section presents three kinds of problem case. Here's the first:

> Pointing at a place, Al says to Bob: "Chris was there."
>
> Pointing at roughly the same place, Bob later reports this by saying: "Al said that Chris was there."

[6] Lewis (1993a, pp.179–80) claims that attending to vagueness places all the candidates in the extension of the relevant sortal *S*. He then argues for the MC approach's mereological interpretation of individuation via the claim that such contexts require a sense in which there's only one *S*. The Problem of the Two undermines this argument: either no sense in which it's true that there's just one *S* is required, or the mereological interpretation of individuation does not provide it.

Schiffer claims the supervaluationist makes Bob's report superfalse if 'there' is vague. He argues as follows. Each sharpening assigns a different precisely delimited region of space to Bob's utterance of 'there'. But Al didn't say of any precise place that Chris was there. So each sharpening makes Bob's report false. So Bob's report is superfalse.

That first case turns on referential vagueness concerning the demonstrative 'there'. The problem generalises. Consider:

> Anna says to Betty: "Chris is bald."
>
> Betty later reports this by saying: "Anna said that Chris is bald."

Different sharpenings of Betty's utterance of 'bald' assign it different precise extensions. But Anna didn't say (anything to the effect) that Chris belongs to any precise extension. So each sharpening makes Betty's report false. So Betty's report is superfalse.

One final case:

> Adama says to Bill: "Baldness is possessed by Chris."
>
> Bill later reports this by saying: "Adama said that baldness is possessed by Chris."

Different sharpenings of Bill's utterance of 'baldness' assign it different precise properties. But Adama didn't say of any precise property that it is possessed by Chris. So each sharpening makes Bill's report false. So Bill's report is superfalse.

Note the use of property-nominalisation to convert predicative vagueness into referential vagueness. If Schiffer's problem is genuine, then supervaluationism faces a problem with referential vagueness, regardless of how it approaches the Problem of the Many.

Although the first and third cases involve referential vagueness, the second doesn't. So why present this as a problem about referential vagueness? One answer is that Schiffer's (1998) paper denies the existence of vague places, despite acknowledging vague properties and propositions. This creates a special problem for the first case, not shared by the other two. Schiffer (2000) retracted this, but a problem peculiar to vague names remains.

Bob and Bill's reports might be re-parsed thus:

> "Al said, of that place there, that it's where Chris was."
>
> "Adama said, of baldness, that it's possessed by Chris."

The truth of *de re* constructions like these is typically insensitive to which expressions are used to denote the referents of the terms in the positions occupied by 'that place there' and 'baldness'.[7] So if these truths become false when terms like 'precise place *p*' and 'precise property *F*' are substituted into those positions, then that must be because Al and Adama's original statements weren't about their referents. But since (i) each sharpening assigns a place or property to 'that place there' and 'baldness' which, let us imagine, is (or could in principle be) designated by terms like those, and (ii) such substitutions do make Bob and Bill's reports false, it follows that (iii) each sharpening makes the reports false.

No similar re-parsing of Betty's utterance to place 'is bald' in a position open to substitution for co-designating (or even analytically coextensive) predicates is possible. Changes in the truth-value of Betty's report when 'belongs to precise extension *e*' is substituted for 'is bald' therefore cannot be attributed to differences in the semantic values of those expressions. We might therefore be unmoved by the second case, despite finding the first and third persuasive.

3.2.1.2 Diagnosis

This section presents an account of the truth-conditions of indirect speech reports and uses it to diagnose the source of Schiffer's complaint.

What are the truth-conditions of indirect speech reports? Well, if Rosie said that grass is green, then there is something Rosie said, namely, that grass is green. This suggests that indirect reports ought to be construed as asserting the obtaining of a relation between a speaker and a potential content, or proposition. The report

[7] An example to illustrate. Suppose Lois Lane has never met Clark Kent, though she is an avid follower of Superman's adventures. Then if you and I know that Kent is Superman, we might report her utterances to one another thus: "Lois said, of Kent, that he saved a child from a falling meteorite yesterday."

will be true iff speaker and proposition really do stand in this relation. Let us denote the proposition that p using '$\langle p \rangle$'. Then the natural truth-condition is:[8]

$\ulcorner S$ said that $A \urcorner$ is true iff:

(i) S uttered a sentence that expressed $\langle p \rangle$; and

(ii) $\ulcorner$that $A\urcorner$ refers to $\langle q \rangle$; and

(iii) $\langle p \rangle = \langle q \rangle$.

This isn't uncontroversial, but it suffices for our purposes. It captures the idea that the goal of an indirect report is to state the content of another's utterance (though not necessarily in the way that they did). An adequate semantics for indirect reports must surely respect this. The apparatus of expressing and referring to propositions merely provides a (natural and plausible) gloss on this.

This truth-condition needs supplementing with accounts of when a sentence expresses a proposition, when a 'that'-clause refers to a proposition, and when $\langle p \rangle = \langle q \rangle$. The next section adds these to the Sharpening View. That framework already provides a model of the association of linguistic items with contents. By conceiving our formal object-language as used to make statements by a community, these additions allow us to use it as a simple model of speech reports also. But before this can provide a satisfactory response to Schiffer, we need to pinpoint the source of his objection.

The problem is a mismatch between (a) the proposition expressed by the sentence being reported, and (b) the propositions that sharpenings assign to the report's 'that'-clause. Recall Schiffer's claim that Al didn't say of any precise place that Chris was there, or equivalently, that Al's utterance didn't express any of the precise singular propositions sharpenings assign to Bob's utterance of 'that Chris was there'. Given the truth-condition above, this amounts to: Al uttered a sentence that expressed $\langle p \rangle$, but there's no sharpening s such that Bob's 'that'-clause s-refers to $\langle p \rangle$. Why not? Schiffer seems to be assuming that Al's utterance expressed a single proposition (not about any precise region of space), while Bob's 'that'-clause

[8] We make two simplifying assumptions. First assumption: sentences are uttered only in order to make statements. Second assumption: sentences are used only literally. Our interest is in the semantics of speech reports, not the analysis of various uses of language.

vaguely refers to many precise propositions.[9] Neither the Supertruth nor Sharpening theorist should grant this assumption.

The Supertruth theorist regards cross-sharpening variation about the referent of a 'that'-clause as representing vagueness about which proposition it picks out. But on their view, each vague language has a unique intended vague semantic structure: sentences aren't vague because it's vague which proposition they express, but because they express vague propositions. The Supertruth theorist should therefore treat the referents of 'that'-clauses as sharpening-invariant: for each sharpening s, $\ulcorner$that $A\urcorner$ s-refers to the vague proposition expressed by A. So no sharpening assigns a precise proposition to Bob's 'that Chris was there'. And on no sharpening does Bob report Al as saying something of any precise place. So Schiffer's objection fails.[10] We won't develop this further because we've rejected the Supertruth View. So let us consider the Sharpening View instead.

3.2.1.3 Cure

We want to extend our Sharpening-theoretic model of vagueness to accommodate indirect reports. If the truth-condition suggested above is correct, then it should hold on every sharpening that respects the intended senses of 'said' and 'that'-clauses. Hence for any such sharpening s:[11]

[9] Note that in order to state the problem, we've had to relativise 'that'-clause reference to sharpenings without similarly relativising the expression of propositions by sentences, or which place Al said something of. The responses canvassed here dissipate the problem by enforcing uniform derelativisation (on behalf of the Supertruth View) or uniform relativisation (on behalf of the Sharpening View).

[10] This might seem to introduce non-uniformity in truth-conditions: the s-referent of $\ulcorner$that $A\urcorner$ isn't the proposition s-expressed by A or determined solely by the s-values of the constituents of A. A paratactic treatment of indirect reports avoids this. On this view, $\ulcorner S$ said that $A\urcorner$ decomposes into two sentences. One is A. The other demonstrates A: S said *that*☞A. We can then regard this demonstrative as indicating A in order to refer to the (vague) proposition it expresses, in much the same way as we refer to colours by indicating objects that possess them. Since A itself isn't a constituent of the second sentence, there's no argument for non-uniformity in truth-conditions. Rumfitt (1993) develops a paratactic proposal along broadly similar lines.

[11] As stated, this truth-condition might be inadequate. It permits supertrue reports that are less vague than the original statement: a report can be supertrue provided any subset of propositions expressed by the original statement get assigned to its 'that'-clause. Two responses are available. (a)

$\ulcorner S$ said that $A\urcorner$ is s-true iff:

(i) S uttered a sentence B such that B s-expresses $\langle p\rangle$; and

(ii) $\ulcorner$that $A\urcorner$ s-refers to $\langle q\rangle$; and

(iii) $\langle p\rangle = \langle q\rangle$.

Think of our formal object-language as used by a community. Then we can take the uttering of sentences as part of our background metatheory, rather than as analysable within the formal framework.[12] Our framework already represents the possible assignments of logically relevant content to linguistic items. We'll add to this an account of when a sentence expresses a proposition (under an assignment), when a 'that-'clause refers to a proposition (under an assignment), and also the identity conditions of propositions. The result will hopefully be a reasonably accurate representation of the semantics of speech reports that can accommodate the clear truth of indirect reports of vague speech.

At the end of §3.2.1.2 we denied, on behalf of the Supertruth theorist, that 'that'-clauses receive different propositions on different sharpenings. The present response differs. We'll vary the proposition expressed by the original sentence in tandem with the proposition assigned to 'that'-clauses used to report it.

When does a sentence express a proposition? Propositions are what speakers say: if S said that A, then there's something S said, namely $\langle A\rangle$. $\langle A\rangle$ is the content of A. Since the (logically relevant) content of a sentence is its truth-condition, we can see proposition-talk as a nominalisation of truth-conditions. The Sharpening View represents the association of sentences with truth-conditions using a recursive definition of $\Vdash$. So where 'p' stands in for a sentence in this definition:

> Suppose the metasemantic facts are as M represents them. Then A s-expresses $\langle p\rangle$ iff: $s, M \Vdash A$ iff p.

Retain the truth-condition on the grounds that such reports are misleading, but not false. (b) Add the following to (i)–(iii): if S uttered a sentence that, for some sharpening s, s-expresses $\langle p\rangle$, then, for some sharpening t, $\ulcorner$that A$\urcorner$ t-refers to $\langle p\rangle$.

[12] For simplicity, we'll homophonically translate names for speakers into the object-language and ignore contextual variation in sentence-content between utterance and report.

It doesn't automatically follow that vague, and hence multiply interpreted, sentences express many propositions because we haven't yet given an account of when $\langle p \rangle = \langle q \rangle$.

When does a 'that'-clause refer to a proposition? The natural answer is that they refer to the proposition expressed by the sentence from which they're formed:

> Suppose the metasemantic facts are as M represents them. Then ⌜that A⌝ s-refers to $\langle p \rangle$ iff: $s, M \Vdash A$ iff p.

Putting these pieces together:

> Suppose the metasemantic facts are as M represents them. Then ⌜S said that A⌝ is s-true iff:
>
> (i) S uttered a sentence B such that: $s, M \Vdash B$ iff p; and
>
> (ii) $s, M \Vdash A$ iff q, and
>
> (iii) $\langle p \rangle = \langle q \rangle$.

Two comments before we continue.

Firstly, this makes the reports in the original cases supertrue, even without an account of proposition-identity. Those cases use the same sentence in the original utterance and report.[13] So the utterance s-expresses the same proposition as the 'that'-clause s-refers to, for any sharpening s. An account of proposition-identity is needed only when different sentences are used in the original statement and report (since we're ignoring context-sensitivity).

Secondly, consider a notion of proposition on which multiply interpreted sentences do express many propositions. Our truth-condition makes it vague what was said if the original statement was vague. For the extension of 'said', as defined by clauses (i)–(iii), varies across sharpenings: if S said that A, then it's s-true that S said only what A s-expresses. We could avoid this by replacing (i) with:

(i′) S uttered a sentence B such that, for some sharpening $t \in M : t, M \Vdash A$ iff p.

[13] The first case is tricky because it features demonstratives in the original utterance and report. We'll address that shortly.

This brings the advantage of permitting a conception of speech-reports as an object-language reflection of metalinguistic claims about intended interpretation.[14] The unmodified account cannot do so because claims about the range of intended interpretations ought to be constant across those interpretations. Luckily, we need not decide between (i) and (i′) here because both make the reports supertrue in each of our three cases. So we'll stick with the simpler condition (i).

Finally, to complete our response, we need an account of the identity conditions of propositions. There are many options. Each seems to capture *a* legitimate notion of content, or of what was said. Instead of arguing about the One True Notion of Proposition, we can allow that different notions might be relevant to the various projects in service of which different reports are made. (Moore, 1999, argues for a similar view.) Different notions of proposition are characterisable using different types of permissible transformation: $\langle p \rangle = \langle q \rangle$ iff permissible transformations convert p into q. Some candidate permissible transformations are:

> The identity transformation: $\langle p \rangle = \langle q \rangle$ iff $p = q$. This very fine-grained notion is maximally sensitive to syntactic/compositional structure: $\langle A \vee B \rangle \neq \langle B \vee A \rangle$ when $A \neq B$.
>
> As above, but also permutation of conjuncts and disjuncts. This notion is less sensitive to compositional structure. Such propositions are roughly akin to Fregean propositions in being individuated by presentations of semantic values.[15]
>
> As above, but also substitution of co-referential terms for elements of the domain. This roughly corresponds to a singular proposition in being insensitive to how objects are designated.
>
> As above, but also substitution of co-referential terms for set-theoretic constructs from the domain. This is akin to a Russellian proposition, a structured complex of objects and properties.

[14] The advantage is an explanation of a theoretical metasemantic concept in familiar terms.

[15] Variants allow substitution and insertion of double negation signs, and interchange of quantifiers for their duals. We'll assume the following notions permit these transformations.

> Interderivability. This corresponds to a conception of propositions as sets of (logically) possible worlds.[16]

Now everything's in place, let's apply it to the first problem case in §3.2.1.1.

Brian Weatherson (2003, §1) suggests the following truth-condition:

> Bob's report 'Al said that Chris was there' is s-true iff: if Al's utterance of 'there' s-refers to a place x, then so does Bob's.

Think of 'Chris was there' as an atomic predication 'F(there)'. Al's statement and Bob's 'that'-clause contain this same predicate, 'Chris was (located at)...'. Do the demonstratives in statement and report count as the same singular term? Let us suppose not; maybe Al and Bob had to point in quite different directions. Then Weatherson's truth-condition for this particular case follows from our more general account, provided that proposition-identity is insensitive to substitution of co-referring terms for elements of the domain. Is Bob's report supertrue? That depends on whether the following penumbral connection holds:

> For every sharpening s, Al's 'there' s-refers to x iff Bob's 'there' s-refers to x.

Since Bob pointed at roughly the same area as Al and intended to use his demonstrative to report Al's statement, it is very plausible that only interpretations that respect this constraint will count as intended: Bob used his demonstrative deferentially to how Al used his. So Bob's report is supertrue. Similar remarks apply to the other two cases. So Schiffer's objection fails.

3.2.1.4 Vague and precise contents

Our conclusion that Schiffer's argument fails may have been too hasty. We assumed that a report is s-true if s assigns the same proposition to its 'that'-clause as to the original statement. It's unlikely this would satisfy Schiffer. He seems to deny that any proposition assigned by any sharpening to a 'that'-clause is also assigned to a vague utterance: recall Schiffer's claim that Bob didn't say of any precise place that Chris was there; his utterance didn't express any precise proposition about

[16] We ignore higher-order logics and incomplete deductive systems for simplicity.

any precise place. But this is just the negation of the Sharpening View. It therefore carries no weight without supporting argument.

An argument is nearby. Sharpenings assign *precise* places to Al's 'there' and *precise* propositions to his 'Chris was there'. This raises two worries. Firstly, it implies that what Al said was precise, even though the sentence he used to say it was vague. Secondly, how can Al's statement be vague if what he said is precise? If these worries are genuine, then the Sharpening View's account of vagueness collapses. This section addresses this worry.

Note first that the Sharpening View analyses vagueness using a combination of semantic and metasemantic concepts. The vague/precise classification therefore primarily applies to content-bearers, not to their contents. We'll use 'presentation' as a neutral term for any kind of content-bearer. Our first task is to extend the vague/precise classification from presentations to contents. Three suggestions follow.[17]

According to the first suggestion:

> x is vague (precise) iff x is/could be the content of some vague (precise) presentation.

On this view, vagueness and precision aren't mutually exclusive classifications of contents.[18] Since vague presentations always have vague contents, what Al said was vague. But this doesn't completely alleviate the problem because what Al said may well also be precise.

According to the second suggestion:

> x is vague (precise) iff x is/could be the content only of vague (precise) presentations (and is/could be the content of some vague (precise) presentation).

The closing parenthetical comment excludes trivially vague and precise contents. On this view, vagueness and precision aren't exhaustive classifications of contents.

[17] A less concessive, though probably sound, response to Schiffer denies that there are any precise contents on the grounds that only presentations can be vague or precise (and then, only in their role as presentations). We won't develop this here.

[18] Should vagueness and precision be mutually exclusive classifications of contents? That depends on how deeply entrenched in our conceptual scheme that incompatibility is, as applied to contents rather than presentations. I'm inclined to think that it's not very deeply entrenched.

Since Al uttered a vague sentence, what he said wasn't precise. But there's no guarantee that it will be vague either. So this suggestion doesn't completely alleviate the problem either.

A more promising suggestion modifies the logical form of attributions of vagueness to contents, by relativising them to presentations:

x is vague (precise) relative to α iff α is vague (precise) and x is a content of α.

On this view, a content can be vague relative to one presentation and precise relative to another. The important notion when assessing the vagueness of what someone said, is whether it was vague relative to the way they said it (whether they conceptualised it as vague). Since Al's 'Chris was there' is vague, so is what he said, relative to the way he said it. And in this same sense, what Al said—i.e. that Chris was there—isn't precise either. On this approach, the propositions that sharpenings assign to Al's original statement and Bob's report aren't precise relative to either the statement or report. The objection from the precision of those propositions and the places they are about, therefore fails.

This provides the resources to respond to another of Schiffer's objections. He tries to commit the supervaluationist to truth-conditional ambiguity in the form 'baldness is...'. Consider:

Baldness is possessed by Chris.
Baldness is a vague property.

Suppose that 'baldness' is the only vague expression here. Then the first sentence is clearly true iff each property each sharpening assigns to 'baldness' is possessed by Chris. Since, we may suppose, Chris does possess each such property, the first is clearly true. And the second sentence is clearly true iff each property each sharpening assigns to 'baldness' is vague. But since, Schiffer claims, each such property is precise, the second is clearly false. Since it should be clearly true, Schiffer claims that its s-truth must turn not on whether the s-referent of 'baldness' belongs to the s-extension of 'is vague', but on whether the word 'baldness' has many intended interpretations.[19] But then there is truth-conditional ambiguity in 'baldness is...':

[19] We treat 'x is a vague property' as analysable into 'x is vague and x is a property'.

the s-truth-conditions of only the first sentence displayed above turn on whether the s-referent of 'baldness' belongs to the s-extension of the expression that replaces the dots.

Schiffer's argument is fallacious. He makes an unwarranted leap from the claim that (a) the s-truth of 'baldness is a vague property' turns on whether 'baldness' has many intended interpretations, to the claim that (b) the s-truth-condition of 'baldness is a vague property' is not that the s-referent of 'baldness' belongs to the s-extension of 'is vague'. The Sharpening theorist grants (a) but may reject (b).

The s-extension of 'is vague relative to' is the class R of pairs $\langle \alpha, x \rangle$ such that α has many intended semantic values, one of which is x. We need to obtain an s-extension for the unrelativised 'is vague' from this. Some means of closing the second argument position of 'is vague relative to' is needed. The natural suggestion is:

> The s-extension of 'is vague', as that predicate occurs in $\ulcorner \alpha$ is vague$\urcorner$, is $\{x : \langle \alpha, x \rangle \in R\}$.

Since each property assigned by any sharpening to 'baldness' is vague relative to 'baldness', this makes it supertrue that baldness is a vague property. On this view, the s-truth of 'baldness is vague' turns on whether 'baldness' has many intended interpretations because the s-extension of 'is vague' does. But since the s-truth-condition of 'baldness is vague' is that the s-referent of 'baldness' belongs to the s-extension of 'is vague', there's no argument for truth-conditional ambiguity here. There is cross-sentence variation in the s-extension of 'is vague'. This brings departure from compositionality: the s-truth-condition of $\ulcorner \alpha$ is vague$\urcorner$ isn't a function of the s-values of α and 'is vague', but also of α itself. But since this brings no loss of systematicity, there's no reason to find it objectionable.[20] Schiffer's objection therefore fails.

[20] Arguments from, e.g., the productivity of language to compositionality only seem to require a finitely axiomatisable means of determining truth-conditions, not narrow compositionality of semantic values.

3.2.2 Barnett on incomplete definitions

David Barnett (2008) argues that supervaluationist accounts of vague reference are incompatible with the following constraint on reference:

Referential Uniqueness (RU): A singular term refers only if features of its use determine, of some unique thing, that that thing is its referent.

The incompatibility is supposed to arise because our use of 'Tibbles' does not determine, of any Tibbles-candidate, that it is the referent of 'Tibbles'; the candidates are all on a par in that respect.

One line of response correlates (i) which object a sharpening s counts as being determined as the referent of 'Tibbles' by our use of 'Tibbles', with (ii) the s-referent of 'Tibbles':

> For each term α and sharpening s, the s-extension of $\ulcorner x$ determines that α refers to $y\urcorner$ is a relation that holds between our use of α and the s-referent of α.

This makes the following supertrue: there is something x such that our use of 'Tibbles' determines that 'Tibbles' refers to x. But since 'Tibbles' refers to different candidates on different sharpenings, there's be no object x of which it's supertrue that our use of 'Tibbles' determines that 'Tibbles' refers to x. The *de re* formulation of RU is intended to block this.

§1.4.7 presented a similar problem for singular thought, as opposed to linguistic reference, taken from Unger (1980, §12*A*). Since no Tibbles-candidate is singled out in preference to any other as the subject of our Tibbles-thoughts, Unger denies that we have any singular thoughts about Tibbles. This problem is partly addressed here, and partly in the next section.

Barnett also endorses a parallel constraint on predication that, by similar reasoning, should be incompatible with supervaluationism:

Predicative Uniqueness (PU): A predicate expresses a property only if features of its use determine, of some unique property, that that property is expressed by the predicate.

PU isn't needed to create trouble for the supervaluationist approach to predication because a property-name formed by nominalising a vague predicate F will lack a uniquely determined referent if F lacks a uniquely determined extension. Predicative vagueness and property-nominalisation alone should suffice for conflict with RU.

Barnett's objection carries weight only against the Supertruth View; for only on that view is no referent determined for a vague name. The next section elaborates the Sharpening theorist's response. We focus primarily on RU, though similar remarks apply to PU and singular thought.

3.2.2.1 Solution

What is it for our use of α to determine, of x, that α refers to x? The answer is: for some intended interpretation s, α s-refers to x. So the Sharpening View implies that our use of 'Tibbles' does determine, of each Tibbles-candidate, that 'Tibbles' refers to it; for 'Tibbles' refers to each under some intended interpretation. Since there's no need to vary what counts as an intended interpretation across those interpretations, it will even be supertrue of each candidate that our use of 'Tibbles' determines that 'Tibbles' refers to that candidate. Hence even the *de re* aspect of RU is unproblematic.

In order to conflict with the Sharpening View, RU must therefore require that use determine a unique intended interpretation. But then RU is just the negation of the Sharpening View. Why should that be a constraint on reference? An argument is required. Yet neither Barnett's argument for RU nor the use to which he puts it, requires uniqueness of intended interpretation. We begin with his uses of RU.

In §2 of his article, Barnett uses RU to argue that the following fail to introduce referring names:

> Let 'Bitz' name a resident of New York.
>
> Let 'Frib' name a five-year-old child in Nigeria.
>
> Let 'Ball#1' refer to one of the two balls in this urn.

Barnett's goal is to undermine arguments from incomplete stipulations to indeterminacy. His argument is, in essence, that these stipulations don't determine any intended interpretation because they fail to determine, of any object x, that the term in question refers to x; for no object x do the stipulations make the state of x relevant to the truth of sentences featuring 'Bitz', 'Frib' or 'Ball#1'. A term that lacks intended interpretation can hardly be a source of indeterminacy of meaning. But it's consistent with this that some uses of names might make more than one interpretation intended, or more than one object relevant to the truth of a sentence (featuring only that one name).

Here's Barnett's argument for RU:

> "The constraint has the air of a truism. By definition a singular term purports to refer to a single thing: if it has a referent, it has a unique referent. And it is a platitude about meaning that words have their semantic features determined solely by features of their use (where use is construed broadly, to include both speaker intentions and relations to their environment). Hence, if a singular term refers, features of its use must determine a unique referent for it. (Do not confuse this constraint with outright rejection of indeterminacy of reference; it does not exclude indeterminacy as to *which* object is so uniquely determined.) We have what appears to be a trivial constraint on reference for singular terms." (Barnett, 2008, p.173)

Construed as an argument for uniqueness of interpretation, this is fallacious.

Consider the first premiss: if a singular term has a referent, then it has a unique referent. This has two disambiguations:

(i) If a singular term α has an intended interpretation, then α has a unique intended interpretation, and that interpretation assigns α a unique referent.

(ii) If a singular term α has an intended interpretation, then that interpretation assigns α a unique referent.

(i) is obviously question-begging in the present context. A persuasive argument can therefore involve only (ii). But when combined with the claim that use alone determines meaning—that use alone determines intended interpretation—(ii) yields only:

> If a singular term α has an intended interpretation s, then use alone determines that s is an intended interpretation of α and s assigns α a unique referent.

This is compatible with a singular term having many intended interpretations. To rule that out, Barnett needs the question-begging (i). He therefore provides no argument for a reading of RU incompatible with the Sharpening theorist's account of vague reference.

3.2.3 McGee and McLaughlin on *de re* belief

De re and *de dicto* readings of 'Ralph believes that Tibbles is a cat' are typically distinguished by quantifying into the scope of 'believes':

> If Ralph believes that Tibbles is a cat, then Ralph's belief is *de re* iff, for some object x, Ralph believes that x is a cat.

McGee and McLaughlin (2000, pp.144-7) argue that this creates a problem for supervaluationist accounts of vague reference.

3.2.3.1 The problem

Suppose Ralph believes that Tibbles is a cat. His belief is (super)true. The extension of 'cat' varies across sharpenings. So there's no object x of which it's supertrue that Ralph believes that x is a cat; for if there were, his belief wouldn't be true on all sharpenings. So Ralph's belief is not *de re*. Since this turned on no features specific to this case: *de re* belief about ordinary objects is impossible (or maybe just extremely unlikely).

There is a subtlety here. Isn't it true under each sharpening s that there's something of which Ralph believes that it is a cat, namely the s-referent of 'Tibbles'? If so, then quantification-in is legitimate: it's supertrue that, for some object x, Ralph believes that x is a cat. Why isn't this sufficient for Ralph's belief to be *de re*?

The reason is that McGee and McLaughlin deny that it's true under any sharpening that there's something Ralph believes to be a cat:

> "[T]here is no obvious way that supervaluation theory is going to help us here. When we examine acceptable models [i.e. sharpenings], we look at different ways of assigning sharp values to the terms of our language. But assigning sharp values to the terms of our language doesn't do anything to sharpen the focus of Ralph's beliefs. If A is an acceptable model, A assigns a unique body of land to 'Kiliminjaro'. But doing this doesn't do anything to answer the question whether Ralph's belief is about A('Kilimanjaro')(+) or A('Kilimanjaro')(-) [where these are two nearly coincident Kilimanjaro-candidates]." (McGee and McLaughlin, 2000, p.146)

McGee and McLaughlin conceive sharpenings as formal representations of classical semantic-structures that depart from the semantic properties of a vague language only by settling all borderline cases. They don't conceive supervaluationism as a theory of vagueness, so much as a formal structure that resembles the structure of vague thought and language in various respects. No sharpening makes it true that there's something Ralph believes to be a cat because sharpenings don't modify, and aren't constituents of, the content of Ralph's beliefs.

McGee and McLaughlin's challenge thus concerns the object-directedness of *de re* belief. In classical semantics, this coincides with quantification-in. Not so if truth is supertruth: it can be supertrue that something is such that..., without it being supertrue of anything that it.... Object-directedness isn't guaranteed by quantification-in alone, but by combining quantification-in with (super)truth:

> If Ralph believes that Tibbles is a cat, then Ralph's belief is object-directed iff it's (super)true of some object x that Ralph believes that x is a cat.

So on the Supertruth View, Ralph's belief isn't object-directed. But is it *de re*? A positive answer requires an account of *de re* belief that doesn't imply object-directedness. We'll leave that to the Supertruth theorists (Weatherson, 2003, §3 makes a start), and turn to the Sharpening View instead.

3.2.3.2 Multiply interpreted *de re* belief

On the Sharpening View, vague content-bearers possess many contents. In particular, vague *de re* beliefs possess many singular contents. When Ralph believes that Tibbles is a cat, he believes each singular classical proposition assigned by any sharpening to 'Tibbles is a cat'. The Sharpening theorist thus claims that Ralph has many similar *de re* beliefs about the Tibbles-candidates.

Two comments. First, sharpenings must assign content to mental states as well as linguistic items: vague thoughts and sentences both express many propositions. Second, each of these singular propositions is object-directed. Ralph's belief therefore permits quantification-in: for many objects x, Ralph believes that x is a cat. McGee and McLaughlin disagree:

> "The possibility that Ralph believes all of the countless singular propositions obtained by supplying Kilimanjaro candidates as arguments of the proposition function [*x is a mountain*] can be readily dismissed, for it implies that, no matter how careful and knowledgeable geographer Ralph may be, his every true belief is accompanied by countless billions of false beliefs." (McGee and McLaughlin, 2000, p.146)

The problem is as follows. McGee and McLaughlin conceive the content of Ralph's cat-beliefs as given by a function f from objects x onto singular propositions such that: $f(x)$ is true iff x is a cat. Let $T_1, \ldots, T_n$ be the Tibbles-candidates. Since there's only one cat on the mat, at most one proposition $f(T_i)$ is true. So, contrary to our claim above, Ralph doesn't believe of each candidate that it is a cat; for if he did, he would believe many false propositions.

The response is that our characterisation of the Sharpening theorist's position above was slightly misleading. On that view, the content of Ralph's singular cat-beliefs isn't given by a single function from objects onto singular propositions, but by a collection of functions $f_1, \ldots, f_n$. Each f_i maps each Tibbles-candidate T_j onto a singular proposition $f_i(T_j)$. For each f_i, exactly one of these propositions is true. And for each Tibbles-candidate T_j, some proposition $f_i(T_j)$ is true. These functions are the sharpenings of 'cat'. The Sharpening theorist claims that Ralph believes each true singular proposition obtained by supplying a Tibbles-candidate to one of these proposition functions. By sharpening the proposition function assigned to 'cat' and the candidate assigned to 'Tibbles' in tandem, Ralph's belief can both have many contents and be supertrue. Ralph doesn't believe that many candidates are cats, but his belief that Tibbles is a cat has many contents, each involving a slightly different way some candidate is believed to be.

McGee and McLaughlin won't permit this. They think of the content of Ralph's cat-beliefs as given by a single proposition function onto vague propositions. But this aspect of the Supertruth View is just what the Sharpening theorist denies. The Sharpening View faces no problem with *de re* belief.

3.2.4 Sorensen on Direct Reference

Roy Sorensen (2000) argues that supervaluationist accounts of vague reference are incompatible with:

Direct Reference (DR): The semantic value of a name is its referent; names contribute only their referents to the truth-conditions of sentences in which they occur.

He describes a scenario in which some explorers introduce 'Acme' to name the first tributary of the river Enigma, which they are about to begin charting.

> "When [explorers] first travel up the river Enigma they finally reach the first pair of river branches. They name one branch 'Sumo' and the other 'Wilt'. Sumo is shorter but more voluminous than Wilt. This makes Sumo and Wilt borderline cases of 'tributary'.... 'Acme' definitely refers to something, even though it is vague whether it refers to Sumo and vague whether it refers to Wilt." (Sorensen, 2000, p.180)

Assume DR: the semantic value of a name is its referent. Then each of 'Acme' and 'Wilt' contributes an object to the truth-conditions of 'Acme is Wilt'. Since the 'is' of identity isn't vague, 'Acme is Wilt' expresses a proposition of either the form $\langle x = x \rangle$ or the form $\langle y = x \rangle$.[21] But such propositions cannot be sharpened. So the supervaluationist technique cannot apply.

[21] Here we adopt Weatherson's (2003) presentation for simplicity. Sorensen's uses the idea that, given DR, names function semantically like variables under an assignment.

The example is poorly chosen. It is questionable whether 'Acme' is semantically directly referential, as opposed to a rigidified description. But we can set this aside; for if any names are directly referential, then names for ordinary objects surely are; in which case, the Lewisian Proposal makes 'Tibbles' relevantly analogous to Sorensen's treatment of 'Acme'.

Only the Supertruth theorist should be moved by this, though not as it stands. They should reject the (classical) conception of propositions Sorensen assumes. Propositions should instead be understood in terms of the supertruth- and superfalsity-conditions assigned them by supervaluationist models. On this approach, an expression's semantic contribution is its role in delimiting the space of sharpenings. How does a name contribute to this? A natural answer is: a name contributes a class of objects, its candidate referents. But this isn't quite right because 'Tibbles' and 'cat' are penumbrally connected. On the Supertruth View, 'Tibbles' makes at least two semantic contributions: (i) its candidate referents; (ii) its penumbral connections to other expressions. Since penumbral connections are analytic connections between meanings, contribution (ii) is incompatible with DR.

On the Sharpening View, DR ought to be relativised to intended interpretations:

Relativised Direct Reference (RDR): For any intended interpretation s, a name's sole contribution to s-truth-conditions is its s-referent.

On one sharpening, 'Acme' contributes Sumo to the truth-conditions of 'Acme is Wilt'. On another, it contributes Wilt. On the first sharpening, 'Acme is Wilt' expresses a proposition of the form $\langle y = x \rangle$. On the second, it expresses a proposition of the form $\langle x = x \rangle$. 'Acme' complies with RDR on both sharpenings, though 'Acme is Wilt' has a different truth-value on each. So RDR is compatible with 'Acme is Wilt' being borderline. The key point is that sentences get supervalued, not the propositions they express. And since penumbral connections are constraints on interpretations, not features of them, they present no threat to RDR.

Weatherson introduces a variant puzzle:

> "[O]ne quite plausible principle about precisifications is that precisifications must not *change* the meaning of a term: they may merely provide referents where none exists. Now the supervaluationist has a prob-

> lem. For it is true that one of ['Acme is Wilt' and 'Acme is Sumo'] is true in virtue of its meaning, since its meaning determines that it expresses a proposition of the form $\langle x = x \rangle$. But each sentence is false on some precisifications, so some precisifications change the meanings of the terms involved.... The best way to respond to this objection is simply to bite the bullet." (Weatherson, 2003, p.498)

The problem is illusory. The Supertruth theorist should deny that the meaning of either 'Acme is Wilt' or 'Acme is Sumo' is a proposition of the form $\langle x = x \rangle$. Hence neither need be true in virtue of its meaning. And the Sharpening theorist should deny that sharpenings provide referents where none exists. Although different sharpenings assign different referents, they don't "fill in the gaps" left by some other reference relation that isn't relativised to sharpenings: s-reference is the only semantic notion of reference. So no bullet-biting is required.

3.3 Three problems with the Problem of the Many

We've seen four objections to supervaluationist accounts of vague reference. The Sharpening theorist should not find them compelling. This leaves us clear to follow the Lewisian Proposal and employ referential vagueness in an analysis of vague boundaries and response to the Problem of the Many. This section presents three problems for this Proposal. §3.3.1 questions whether it provides a genuine solution to the problem. §3.3.2 argues that it cannot accommodate vagueness in the boundaries of self-referrers. And §3.3.3 argues that it brings objectionable commitments in the metaphysics and semantics of time and modality. We ought to seek an alternative.

3.3.1 A genuine solution?

Is the Lewisian Proposal a genuine solution to the Problem of the Many? There are good, but inconclusive, reasons to think not. This section approaches them via a dilemma due to Neil McKinnon (2002). §3.3.1.1 presents the first horn, on which sharpenings are *principled*: objects that satisfy the same predicate share some (not overly disjunctive) property that suffices for satisfying that predicate. §3.3.1.2

presents the second horn: sharpenings are unprincipled. McKinnon argues that neither horn is acceptable. Even if he is wrong, only the second horn is tenable. §3.3.1.3 shows that this commits the supervaluationist to the extrinsicality of the property(s) expressed by 'cat'. §3.3.1.4 suggests that this undermines the claim that the present approach provides a genuine solution.

3.3.1.1 McKinnon's first horn: principled sharpenings

Suppose there are two cats on the mat, Tibbles and Sophie. The following should be clearly true:

There are exactly two cats on the mat.

So for each sharpening s, exactly one Tibbles-candidate and one Sophie-candidate s-satisfy 'cat'. Suppose that sharpenings are principled. Then if x and y both s-satisfy 'cat', then they share some property sufficient for s-satisfaction of 'cat'. But any relevant property shared by a Tibbles-candidate and a Sophie-candidate will also be shared by all the Tibbles-candidates. For the Tibbles-candidates are much more like one another in cat-respects than they are any Sophie-candidate. (Perhaps Tibbles and Sophie belong to different breeds.) So all the Tibbles-candidates s-satisfy 'cat'. Since s was arbitrary: it's superfalse that there are exactly two cats on the mat.

The defender of principled sharpenings has two options. The first is simply to hope that there are suitable properties to provide enough principled sharpenings. Such wishful thinking should be given no credence. The second is to present enough examples to make it plausible that there are sufficient properties to provide enough principled sharpenings to accommodate all of Tibbles and Sophie's evident vagueness. McKinnon considers and rejects several candidates, e.g.: specific shapes and ratios of interior to exterior densities. The candidates seem either so specific as to prevent co-satisfaction of 'cat' by a Tibbles-candidate and a Sophie-candidate, or there's no guarantee that they'll be possessed by exactly one candidate from each collection. So let us set principled sharpenings aside.

3.3.1.2 McKinnon's second horn: arbitrary sharpenings

McKinnon (2002, §3) endorses:

Non-arbitrary differences (NAD) For any cat and non-cat, there is a principled difference between them in virtue of which the one is a cat and the other not.

Non-arbitrary similarities (NAS) For any two cats, there is a principled similarity between them in virtue of which they are both cats.

These induce penumbral connections between 'Tibbles' and 'Sophie' that McKinnon thinks spell trouble for the present approach.

Suppose that Tibbles-candidate T^+ s-satisfies 'cat', while Tibbles-candidate T^- does not. Then by NAD: there is a principled difference between T^+ and T^- in virtue of which this is so. But T^+ resembles T^- much more closely in cat-respects than either does any Sophie-candidate. So each principled difference between T^+ and T^- is also a principled difference between T^+ and each Sophie-candidate. So no Sophie-candidate s-satisfies 'cat'.[22] Since s was arbitrary: it's superfalse that there are exactly two cats on the mat.

Now suppose that Tibbles-candidate T^+ and Sophie-candidate S both s-satisfy 'cat'. Then by NAS: there is a principled similarity between T^+ and S in virtue of which this is so. But T^+ resembles T^- much more closely in cat-respects than it does any Sophie-candidate. So each principled similarity between T^+ and S is also a principled similarity between T^+ and T^-. So T^- also s-satisfies 'cat'. Since s was arbitrary: it's superfalse that there are exactly two cats on the mat.

3.3.1.3 Extensive overlap and intrinsicality

Weatherson (2003, §5) responds to the second horn by, in effect, varying what counts as a principled difference across sharpenings. He begins with the equivalence:

> x is a cat iff x is a cat-candidate that does not extensively overlap any cat (other than x).

[22] This assumes that if something is an F in virtue of being a G, then being a G suffices for being an F.

This is true on all sharpenings that place exactly one Tibbles-candidate and one Sophie-candidate into the extension of 'cat'. It's true on any such sharpening that any two cats resemble one another in respect of their not extensively overlapping any cat other than themselves. Such sharpenings also make it true that any cat and any cat-candidate that fails to be a cat differ in that same respect. Hence, Weatherson claims, NAD and NAS are supertrue. There are two problems.

The first problem is that NAD and NAS don't involve material biconditionals, but an 'in virtue of' locution. They don't simply assert the existence of principled similarities and differences, but claim that similarities and differences in respect of being cats obtain in virtue of these other principled similarities and differences. But the Problem of the Many arises because the condition *x is a cat-candidate that does not overlap any cat (other than itself)* determines a unique Tibbles-candidate only given a prior selection of some unique candidate as a cat. So on Weatherson's proposal, similarities and differences in respect of being a cat-candidate that does not extensively overlap any cat (other than itself) obtain in virtue of similarities and differences in respect of being a cat, rather than, as NAD and NAS require, *vice versa*. Weatherson has two responses available.

Firstly, he may point out that 'in virtue of' locutions are notoriously murky. It's far from obvious what their content amounts to, or what constrains their correct usage. If these doubts are well-founded, then only the weaker biconditional readings of NAD and NAS may legitimately be insisted on. Secondly, he may (and does) claim that the stronger reading amounts to a demand for a (possibly partial) analysis of 'cat'. There is no reason to expect that this will be possible, especially for an apparent natural kind term like 'cat'. So let us turn to the second problem for Weatherson's proposal.

Weatherson's proposal brings cross-sharpening variation in the respects of resemblance and difference that satisfy NAD and NAS. On sharpening s, cats resemble one another in respect of not extensively overlapping any cat-candidate to which 'cat' s-applies. And on sharpening t, cats resemble one another in respect of not extensively overlapping any cat-candidate to which 'cat' t-applies. Since s and t assign different extensions to 'cat', these are different respects of resemblance. But such similarities and differences obtain only because of the existence of the semantic structures s and t. They are no more genuine, principled or intrinsic resemblances and differences than x and y's resemblance in respect of belonging to $\{x, y\}$, or their difference in respect of belonging to $\{x\}$. If 'cat' distinguishes between objects on such slim grounds, then it does not mark a genuine, principled or intrinsic distinction. The only such distinction in the vicinity is that between the cat-candidates and everything else. But that's not marked by 'cat' on any sharpening.

So if sharpenings are unprincipled, then 'cat' does not mark a natural kind or intrinsic property, regardless of whether McKinnon is right to insist on NAD and NAS. Is this objectionable? One reason to think not appeals to the apparent maximality of 'cat'. But as §1.1.4.2 showed, the argument from maximality to non-intrinsicality assumes that maximality is a semantic feature of predicates, rather than a metaphysical feature of the boundaries of objects. This is tantamount to assuming that the property expressed by 'cat' is extrinsic. Maximality therefore does not provide independent or theory-neutral reason to deny that the cats form a natural kind.

Regardless of whether we ought to deny that the cats form a natural kind, advocates of the present must do so. This raises serious doubts about whether it provides a genuine solution to the Problem of the Many.

3.3.1.4 Genuine solution or semantic trickery?

§1.4 showed that the Problem of the Many generates conflict with:

> There is not widespread causal overdetermination of the effects of ordinary objects by ordinary objects.
>
> Your actions are free, in the sense of not being entailed by those of any person distinct from you.
>
> You make real choices, independent of those made by any other conscious being.
>
> It is an undoubtable Moorean fact that I'm the only person in my chair.
>
> Our ordinary moral judgements are not in radical error.

> You are the only thinking and experiencing conscious being in your chair.

Penumbral constraints like the following resolve this conflict by making the claims above supertrue:

> The *s*-extension of 'ordinary object' is the union of the *s*-extensions of all ordinary sortals, 'cat', 'human', 'dog',. . . .
>
> The *s*-referent of 'you' is the only one of your person-candidates in the *s*-extensions of 'person', 'conscious', 'experiencer', 'chooser' and 'thinker'.
>
> The *s*-referent of 'I' is the only one of my person-candidates in the *s*-extension of 'person'.
>
> The *s*-extension of, for example, 'murderer' is a subset of the *s*-extension of 'person'.

Does this semantic technique solve the initial problems? That depends on what those problems are.

On one account, the Problem of the Many is a problem because it seems to show that certain sentences possess truth-values which they ought not to. If this is right, then we have a genuine solution because the supervaluationist technique ensures a proper distribution of truth-values (though §3.3.2 and the end of the present section question even this). But this is not the only account.

On an alternative account, the Problem of the Many is problematic because it implies an overabundance of certain kinds of object. The preceding section showed that the supervaluationist technique does not address this. Although it's not true on the present approach that my person-candidates are all conscious, they are intrinsically just like conscious beings; let us say that they are conscious*. If there's any important or intrinsic distinction amongst objects in the vicinity of consciousness, it's that between the conscious* things and the rest.[23] Yet I'm not the only conscious* being in my chair; my actions and choices are entailed by the actions* and choices* of many other conscious* beings. Is this problematic? If so, then the

[23] An important distinction amongst objects contrasts with a distinction amongst objects in respect of their relationships to potential semantic structures and our use of language.

Lewisian Proposal merely disguises the problem, rather than addressing it. In effect, the *-ed properties provide a significant sense in which it's supertrue that, say, many cats are on Tibbles's mat: there are many cat*'s on Tibbles's mat. It must be shown that this is unproblematic before we can regard the Lewisian Proposal as a genuine solution.

Now, 'conscious*' is a theoretical term, not part of ordinary vocabulary. So one might deny that we're entitled to any intuitive judgements about whether our near-coincidence with many conscious* beings is problematic. In which case, *-ed analogues of the problems above carry no intuitive force. This may be right. But it doesn't show that those *-ed analogues aren't problems. It shows at most that we aren't entitled to a view either way; in which case, we aren't entitled to a view about the success of the Lewisian Proposal either. Without a proper investigation of the metaphysics of causation, action, free will, choice, consciousness, personhood and morality, endorsing that Proposal amounts to simply closing one's eyes and hoping for the best. Furthermore, there is reason to think that we are entitled to just the same intuitive judgements about *-ed concepts, as we are their *-less counterparts.

Are our beliefs about cats beliefs about a certain kind of object, or about the truth-values of sentences? Surely not the latter. Yet the Lewisian Proposal implies that 'cat' and 'conscious' don't mark genuine kinds. Cat*s and conscious* beings are just like cats and conscious beings in every respect that's relevant to the justification of our beliefs about cats and conscious beings. So the justification for our beliefs about cats and conscious beings extends to justify the same beliefs about cat*s and conscious* beings; in which case, we are justified in regarding the *-ed variants of the problems above as genuine problems, if we are justified in so regarding the originals. If this is right, then the Lewisian Proposal doesn't even assign the right truth-values to sentences, never mind resolve any underlying metaphysical problems.

Although this last argument certainly isn't beyond reproach, a proper investigation lies beyond the scope of this thesis. But it does suggest that the burden of proof lies with the Lewisian to show that their "solution" is genuine. It also exposes the Proposal's hidden assumptions about an array of core philosophical disputes. We ought not to endorse it without having undertaken a proper investigation of those issues.

3.3.2 Hawthorne on self-reference

A variant on an argument of John Hawthorne's makes trouble for the One Cat solution (Hawthorne, 2006a). Hawthorne's argument targeted the following combination of views:

> Vague languages satisfy classical logic and semantics, and have a unique intended interpretation.
>
> Vague expressions are *semantically plastic*, i.e. sensitive to indiscriminably slight variations in use. (This was used to justify the idea that the languages of different speakers might have slightly different intended interpretations.)
>
> Vague boundaries involve vagueness about which object's boundaries are at issue.

Our variant differs from Hawthorne's by (i) not requiring a unique intended interpretation, (ii) not assuming semantic plasticity, and (iii) not assuming that the languages of different speakers can have slightly different intended interpretations. Unlike Hawthorne's argument, ours will apply to ourselves as well as to other speakers.

§3.3.2.1 begins with three versions of the argument. §§3.3.2.2–3.3.2.5 consider and reject four responses.

3.3.2.1 The argument

This section presents three versions of Hawthorne's argument.

Version one: Paula Suppose that Paula is sitting on a chair. Suppose also that her boundaries are vague such that there are two Paula-candidates, P^+ and P^-. Since Paula is a person, and clearly the only person sitting on her chair, P^+ and P^- are also the only person-candidates. So there are two sharpening s^+, s^- of our and Paula's shared vague language:

> 'Paula' s^+-refers to P^+.

'Person' s^+-applies to P^+.

'Paula' s^--refers to P^-.

'Person' s^--applies to P^-.

Given this, the argument rest on four seemingly obvious truths:

(2) Paula is the only object in her chair that can speak and think.

(3) Utterances of 'I' by Paula refer to Paula.

(4) Utterances of the form 'a is F' by Paula are true iff 'F' applies to the referent of 'a', as Paula uses those expressions.

(5) Our and Paula's language is vague: s^+ and s^- are its intended interpretations.

Here's how the argument goes.

Suppose it's clearly true that Paula says "I am a person". Each of (2)–(5) are also clearly true. So each is true on each sharpening. We want to know: is it clearly true that Paula's utterance was clearly true? Is it true on each sharpening that Paula's utterance is true on each sharpening? By (5): s^+ is a sharpening of our language. So is it true on s^+ that Paula's utterance was true on each sharpening? The following argument suggests that it isn't, and hence that it isn't clearly true that Paula's utterance was clearly true.

On s^+, 'Paula' refers to P^+. By (2): Paula/P^+, and nobody else, said 'I am a person'. By (5): s^- is a sharpening of Paula/P^+'s language. By (4): Paula/P^+'s utterance of 'I am a person' is true on s^- iff 'person' s^--applies to the referent of 'I', as Paula/P^+ uses 'I'. By (3): 'I' refers to Paula/P^+, as Paula/P^+ uses 'I'. But 'person' doesn't s^--apply to P^+. So it isn't true on s^+ that Paula/P^+'s utterance of 'I am a person' is true on s^-. So it isn't true on s^+ that Paula's utterance was supertrue. So it isn't clearly true that Paula's utterance of 'I am a person' was clearly true. In fact, it's true on each sharpening that Paula's utterance is true on some but not all sharpenings. So it's clearly true that Paula's utterance was borderline. Yet surely it shouldn't be.

Hawthorne's argument delivers this same conclusion. Unlike Hawthorne's argument however, ours applies to ourselves as well as to Paula.

Version two: me Suppose my boundaries are vague such that there are two Nick-candidates N^+, N^- in my chair. Since I am a person and clearly the only person in my chair, N^+ and N^- are also the only person-candidates. So there are two sharpenings of 'person' in my language:

'Person' s^+-applies to N^+.

'Person' s^--applies to N^-.

Five seemingly obvious truths:

(6) I am the person in my chair.

(7) Only the person in my chair can speak and think.

(8) Utterances of 'I' by the person in my chair refer to that person.

(9) Utterances of the form 'a is F' by the person in my chair are true iff 'F' applies to the referent of 'a', as the person in my chair uses those expressions.

(10) My/the person in my chair's language is vague: s^+ and s^- are its intended interpretations.

Suppose it's clearly true that I say "I am a person". Each of (6)–(10) are clearly true. So each is true on each sharpening. Is it clearly true that my utterance was clearly true? Is it true on each sharpening of my language that my utterance was true on each sharpening? By (10): s^+ is a sharpening of my language. So is it true on s^+ that my utterance was true on each sharpening? The following argument suggests that it isn't, and hence that it isn't clearly true that my utterance was clearly true.

On s^+, 'person' applies to N^+. Hence by (6) and (7): the person in my chair/N^+, and nobody else, said 'I am a person'.[24] By (10): s^- is a sharpening of the person in my chair/N^+'s language. By (9): the person in my chair/N^+'s utterance of 'I am a person' is true on s^- iff 'person' s^--applies to the referent of 'I', as the person in my chair/N^+ uses 'I'. By (8): 'I' refers to N^+, as the person in my chair/N^+ uses 'I'. But 'person' doesn't s^--apply to N^+. So it isn't true on s^+ that the person in

[24] The extra premiss (6) is needed to fix a referent for 'I'. It ensures that its s^+-referent is the unique s^+-satisfier of 'person', namely N^+.

my chair/N^+/my utterance of ‘I am a person’ was true on s^-. So it isn’t true on s^+ that my utterance was supertrue. So it isn’t clearly true that my utterance of ‘I am a person’ was clearly true. In fact, it’s true on each sharpening that my utterance of ‘I am a person’ is true on some but not all sharpenings. So it’s clearly true that my utterance was borderline. Surely it shouldn’t be.

It might seem that this relies on ‘person’ applying to only one of my person-candidates on each sharpening, and hence that the Many Cat approach avoids it. The final variant shows that this is mistaken.

Version three: Nick Stick with the same Nick/person-candidates as before. Two more facts about s^+ and s^-:

> ‘Nick’ s^+-refers to N^+.
>
> ‘Nick’ s^--refers to N^-.

Five seemingly obvious truths:

(11) I am the person in my chair.

(12) Only the person in my chair can speak and think.

(13) Utterances of ‘I’ by the person in my chair refer to that person.

(14) Utterances of the form ‘*a* is *b*’ by the person in my chair are true iff the referent of ‘*a*’ is identical to the referent of ‘*b*’, as the person in my chair uses those expressions.

(15) My/the person in my chair’s language is vague: s^+ and s^- are its intended interpretations.

Suppose it’s clearly true that I say “I am Nick”. This and each of (11)–(15) are clearly true, hence true on each sharpening. Is it clearly true that my utterance was clearly true? Is it true on each sharpening of my language that that my utterance was true on each sharpening? The following argument suggests that it isn’t, and hence that it isn’t clearly true that my utterance was clearly true.

On s^+, ‘person’ applies only to N^+. Hence by (11) and (12): the person in my chair/N^+, and nobody else, said ‘I am Nick’. By (15): s^- is a sharpening of the person in my chair/N^+'s language. By (14): the person in my chair/N^+'s utterance of 'I am Nick' is true on s^- iff the s^--referent of 'Nick' is identical to the referent of 'I', as the person in my chair/N^+ uses 'I'.[25] By (13): 'I' refers to N^+, as the person in my chair/N^+ uses 'I'. But 'Nick' doesn't s^--refer to N^+. So it isn't true on s^+ that the person in my chair/N^+'s utterance of 'I am Nick' was true on s^-. So it isn't true on s^+ that my utterance was supertrue. So it isn't clearly true that my utterance of 'I am Nick' was clearly true. In fact, it's true on each sharpening that my utterance was true on some but not all sharpenings. So it's clearly true that my utterance of 'I am Nick' was borderline. So who am I?

Since the Many Cat approach makes no difference to the interpretation of names, it doesn't undermine this version of the argument. One might respond by pointing out that (14) interprets individuative apparatus using identity, when the MC approach interprets it as near-coincidence.

This response fails because there's a variant of the Problem of the Two on which my leg is being amputated (under powerful local anaesthetic). Halfway through the operation, I say "I am Nick". But my body taken as excluding my leg (N^-) and my body taken as including my leg (N^+) aren't nearly coincident. So the referent of 'I', as the person in my chair/N^+ uses 'I', isn't nearly coincident with the s^--referent of 'Nick'. So it still isn't true on s^+ that my utterance of 'I am Nick' was true on s^-.

What responses are available? Let's focus on the first version for simplicity. Since it rests on four premisses, there are four options. None is attractive.

3.3.2.2 First response: deny (2)

This response denies that Paula is the only object in her chair that can think and speak. This alone won't resolve the problem. We assumed that it's clearly true that Paula said "I am a person". So the argument still shows that her utterance wasn't clearly true; but it also shows that several other speaker's utterances of 'I am a person' weren't clearly true too. To succeed, this response must deny that it's clearly true that Paula spoke at all. This is a significant cost. The analogous response to

[25] Although 'I am Nick' isn't of the form '*a* is *b*', the difference is merely typographical.

the second and third versions of the problem implies that it's not clearly true that it was me who spoke. So who did?

Even this doesn't solve the problem however. For whoever spoke, their language is vague such that, on at least one intended interpretation of their language, 'person' doesn't apply to them.[26]

Even if this response could make Paula's utterance clearly true, there would be sharpenings of her language on which it's true that her actions are entailed by those of someone else, specifically the person in her chair. So she wouldn't clearly act freely. In light of these problems, we should reject this response.

3.3.2.3 Second response: deny (3)

This response denies that Paula clearly uses 'I' as a device of self-reference. Rather, she uses it to refer to the object in her chair that she counts as a person: her utterances of 'I' s^+-refer to P^+, and s^--refer to P^-. So it's true on s^+ that Paula uses 'I' to s^--refer to an object in the s^--extension of 'person'. So it is true on s^+ that her utterance of 'I am a person' was true on s^-. Generalising: it's true on each sharpening that her utterance of 'I am a person' was true on each sharpening.

Hawthorne captures the oddness of this view nicely:

> "There is something exceedingly strange about a view according to which ... many people (perhaps most people) do not [clearly] have linguistic devices of self-reference. Relatedly, it is extremely natural to think that if a pronominal device has the conceptual role of the first-person pronoun in a person's cognitive life, then that pronoun will be a device of self-reference." (Hawthorne, 2006a, p.190)

Maybe we can learn to live with this. But the analogous responses to the second and third arguments are even worse: the person in my chair doesn't clearly self-refer using 'I'. But that person is me. I certainly find this hard to believe, and I

[26] Here we have another difference from Hawthorne's puzzle: denying (2) blocks his argument but not ours. The source of the difference is that, on the Sharpening View but not on epistemicism, speakers aren't one-one correlated with intended interpretations of their language. In particular, they aren't one-one correlated with intended referents for their own name.

suspect that you do too for the corresponding problem for your use of 'I'. We ought not endorse so radical a view unless there really is no alternative.

3.3.2.4 **Third response: deny** (4)

This response denies that Paula's utterances of the form '*a* is *F*' are true iff '*F*' applies to the referent of '*a*', as Paula uses those expressions. Rather, Paula's utterances of that form are true iff '*F*' applies to the referent of '*a*', as we use '*F*'and '*a*'.[27] Then it's true on s^+ that Paula's utterance of 'I am a person' is true on s^-; for as we use 'person' (on s^+), it applies to the object to which Paula/P^+ refers using 'I', namely P^+.

This violates a foundational principle of good translation:

> A correct translation of another's utterances assigns them the truth-conditions that their utterer used them to express.

By denying this principle, the present response amounts to an unacceptable form of semantic chauvinism. It's akin to denying the meaningfulness of French on the grounds that English speakers do not use French vocabulary meaningfully. Why should the near-coincidence of the person-candidates in Paula's chair makes a difference to this? We should reject this response.

3.3.2.5 **Fourth response: deny** (5)

This response denies that both s^+ and s^- are intended interpretations of Paula's language. This will need finessing; for it is a datum that Paula's language is vague. The most promising strategy will vary the intended interpretations of Paula's language across intended interpretations of ours: s^+ doesn't count s^- as an intended interpretation. Then since it's true on s^+ that Paula's utterance of 'I am a person' is true on s^+, it's also true on s^+ that Paula's utterance was clearly true. Likewise for s^-. So it's clearly true that Paula's utterance was clearly true.

[27] Care is needed here. For as we use 'I', it refers to ourselves, not to Paula. An exception is needed for 'I': it's Paula's use that matters, not ours. Since (3) ensures that how Paula uses 'I' is determined by how we use 'Paula', this gives the result we need. But this non-uniformity of truth-conditions should already make us suspicious.

This view makes it true on each interpretation that only one Paula-candidate satisfies 'person' on any intended interpretation of her language. But Paula's language is our language. So it's true on each interpretation of our language that only one Paula-candidate satisfies 'person' on any intended interpretation of our language. Likewise for all other people. And applying the analogous response to the second and third variant puzzles: it's true on each interpretation of my language that only one Nick-candidate satisfies 'person' on any intended interpretation of my language. It follows that 'person' and 'Nick' aren't vague. Since Nick is the only person in my chair, and Paula is the only person in her chair, and. . . , it follows that names for persons are non-vague too, and hence that persons (and self-referrers generally) cannot have vague boundaries.

If persons cannot have vague boundaries, why do they appear to? We must either (i) explain away the mistaken appearance of vagueness, or (ii) adopt a non-supervaluationist account of vagueness in the boundaries of persons. Both options are problematic: why not apply this alternative account to all (apparent) vagueness? The supervaluationist should reject this approach to vague boundaries.

None of these responses is satisfactory. The Sharpening theorist therefore ought not to treat vagueness in the boundaries of speakers and thinkers as vagueness about which object speaks and thinks. In the interests of uniformity, they ought not to treat any vague boundaries in this way.

3.3.3 Coincidence, time and modality

§1.4.1 presented several puzzles about the interaction of the Problem of the Many with time and modality. They arose because Tibbles's boundaries (i) become more and less vague over time, and (ii) could have been more or less vague than they actually are. This section argues that in order to solve these puzzles, advocates of the Lewisian Proposal must endorse objectionable theses about the metaphysics and semantics of temporal and modal discourse (§3.3.3.1). §3.3.3.2 explains why these commitments are objectionable. Lewis himself held (versions of) these views: the Lewisian Proposal about vague boundaries is inextricably linked with the rest of his metaphysical framework.

3.3.3.1 Coincidence, persistence and modality

This section argues from the Lewisian Proposal about vague boundaries to the disjunction of counterpart-theory with perdurance-theory. We've already rejected the former (§1.1.2.1). We reject the latter in §3.3.3.2.

Consider the following world w in which Tibbles remains on the mat throughout his life. Tibbles's boundaries are precise until t_1, when hair h becomes his only borderline part: Tibbles-candidates T^+ and T^- nearly coincide after t_1. Because the following is clearly true, T^+ and T^- were on the mat before t_1, when they were coincident:

Tibbles was on the mat before t_1.

Let v be a world just like w until t_1, when Tibbles is annihilated. Tibbles's boundaries are always precise in v. How many cat-candidates are on the mat in v? §1.4.1 argued that no answer is satisfactory.

Suppose there is only one cat-candidate on the mat in v. Four problems arise:

(a) Is the v-cat-candidate T^+ or T^-? Either answer seems arbitrary. This arbitrariness may be relieved by postulating a world u just like v except for which Tibbles-candidate it contains. Since this is the only difference between v and u, it follows that the events leading up to Tibbles's birth were insufficient to bring him into being.

(b) Suppose that v is actual. Then if Tibbles's future boundaries had been more vague than they actually are, then there would have been more cat-candidates on the mat before t_1 than there actually are. So the number of past objects seems to depend on how the future unfolds.

(c) There is a dynamical sense of possibility in which what's possible changes over time: it used to be possible that I wouldn't finish this thesis, but no longer is. If there's only one candidate in v, then it's not possible in v for Tibbles's boundaries to become vague; for there's only one interpretation on which any possible future candidate was on the mat before t_1 and hence no possible future in which there's more than one intended referent for 'Tibbles'. So it's dynamically impossible in v for Tibbles's boundaries to become vague. More generally, it's dynamically impossible in any world for Tibbles's boundaries to become more vague than they ever are.

(d) In w, it's not clearly true that Tibbles would have existed had the boundaries of the cat on the mat been less vague than they are. For in one of the closest worlds to w in which the cat on the mat's boundaries are precise, namely v, one candidate doesn't exist. So on any sharpening that assigns this candidate to Tibbles, it isn't true that Tibbles exists in all the closest worlds in which the cat on the mat's boundaries are less vague than they are in w.

Now suppose that both cat-candidates are on the mat in v, where they permanently coincide despite belonging to the same kind. Two further problems:

(e) There's no reason for both candidates to exist in v, other than that Tibbles's boundaries could have, but didn't, later become vague. The number of objects seems to depend on how things could have been.

(f) There need to be enough candidates on the mat in v to accommodate any possible vagueness in Tibbles's boundaries. So there will be many candidates indeed, even in worlds where Tibbles's boundaries are always precise.

Although I know of no responses to (e) and (f), there seem to be two responses to (a)–(d) capable of granting that there's only one candidate on the mat in v. The first is counterpart-theory. The second is perdurance-theory.

Counterpart-theory decouples the modal and temporal profiles of ordinary objects from questions about their identity across times and worlds. The puzzles all then concern the number of ways of selecting non-present/other-worldly counterparts, rather than the existence of candidates; questions about the number of candidates in a world can then be separated from questions about tracking them across worlds. But since we rejected counterpart-theory in §1.1.2.2, we won't examine this response in detail. Instead, we proceed directly to perdurance.

3.3.3.2 Perdurance

Perdurance is the version of temporal-parts theory that identifies ordinary persisting objects, the referents of ordinary names, with sums of short-lived objects:[28]

> If an object x exists at a time t, then there is something y such that (i) y exists at and only at t, (ii) y overlaps at t exactly those things that overlap x at t, and (iii) y is part of x. We say that y is x's t-part.

On this view, cat-candidates are spatiotemporally extended "worms". The v-candidate is neither T^+ nor T^-, but the restriction of those two worms to times before t_1, when Tibbles's boundaries became unclear.[29] From this perspective, variation in the extent of unclarity in Tibbles's boundaries looks just like Lewis's (1976) account of fission and fusion: fusion occurs when worms that used not to share temporal-parts come to do so; fission occurs when worms that used to share temporal-parts cease to do so. The difference is simply a matter of scale. Let's apply this to problems (a)–(d) before we reject it.

The perdurantist solution In response to (a): there's no arbitrariness about which candidate exists in v because neither does. Instead, the restriction of T^+ and of T^- to times before t_1, when they shared the same stages, exists in v. This restriction is composed by their pre-t_1-parts, each of which exists in v.

The response to (b) is a little more concessive. The number of past objects does depend on how the future turns out, but only in a familiar and (supposedly) unproblematic way. A spatial analogy is helpful.

The number of roads in a town depends on the world outside the town. What appears within a town to be a single road might really be two roads that share a (spatial) segment within the town and diverge outside of it. The number of roads in a town depends on what happens elsewhere because roads can share some of

[28] There are alternative formulations, but this captures the view's core. For discussion see Hawthorne (2006d) and Sider (2001a).

[29] We make the natural simplifying assumptions that (i) temporal parts have their temporal location essentially, and (ii) spatiotemporal worms have their temporal parts essentially. It follows that temporal extent is an essential property of spatiotemporal worms. Only our manner of presentation turns on this.

their spatial segments without sharing all of their spatial segments. The number of roads within a region r thus depends on the relations between road-segments entirely contained within r and road-segments elsewhere. But the number of road-segments entirely contained within r is independent of their relations to road-segments elsewhere. These features of roads are commonplace and unproblematic. The perdurance-theorist appeals to analogous temporal claims in response to problem (b).

The number of cat-candidates (worms) on the mat at a time depends on what happens at other times. What appears at one time t to be a single candidate could really be two candidates that share (temporal) t-segments, but not later t'-segments. The number of candidates at t thus depends on the relations between candidate-segments that exist only at t and candidate-segments that exist only at other times. But the number of candidate-segments that exist only at t is independent of their relations to candidate-segments elsewhen. The number of past candidates therefore depends on the number of future candidates only in the same unproblematic way in which the number of roads within a town depends on what happens outside the town.[30]

As it stands, this solution to (a) and (b) is not fully general because it cannot accommodate (c) and (d). Consider (c). Let T be the candidate-worm in v: T is the only intended referent of 'Tibbles' in v. For any dynamical future possibility u and object x in u, either x is identical to T or it isn't. If it isn't, then, in v, no intended interpretation assigns x in u to 'Tibbles'. So, in v, there's at most one intended referent for 'Tibbles' in any future dynamical possibility. So if Tibbles's boundaries ever aren't vague, then they can't become vague. Similarly for (d): at most one w-candidate is identical to the v-candidate; so if each w-candidate is an intended referent for 'Tibbles' in w, then it's not clearly true in w that Tibbles is on the mat in v.

[30] Does this really solve the problem? Perdurantism implies a formal analogy between the dependence of (i) the number of past objects on the future, and (ii) the number of local objects on the non-local. But it doesn't follow that since (ii) is unproblematic, so is (i). It needs to be shown that the analogy is more than formal, that (i) and (ii) are analogous in respect of being unproblematic. This is just the point at issue.

A modal analogue of perdurance is required:

> If an object x exists at a time t in a world w, then there is something y such that (i) y exists at, and only at, t in w, and (ii) y overlaps at t in w exactly those things that overlap x at t in w, and (iii) y is part of x.[31]

Distinguish a modal-cat-stage wholly contained within v from a "modally extended" cat-candidate. Cat-candidates are fusions of the temporally extended cat-worms wholly contained within worlds. The existence of only a single such worm in v therefore does not reduce the number of cat-candidates (partly) in v. These candidates provide many intended referents for 'Tibbles' both before t_1 in v, and in any dynamically possible future. This resolves (c). It also resolves (d) because the candidate referents for 'Tibbles' in w are cross-world fusions which overlap objects on the mat (that exist only) in v. Hence those candidates exist in v.

Modal perdurance seems committed to Lewis's ontology of worlds as spatiotemporally and causally isolated concrete spacetimes. This is a very high cost; we ought to reject modal perdurance unless we are Lewisian realists about possible worlds. But without modal perdurance, and hence also Lewis's modal ontology, perdurance-theory is not a fully general solution to puzzles (a)–(d). However, there are good reasons to reject perdurance even if we are prepared to accept Lewis's modal ontology.

Perdurance faces two major difficulties. The first is that it brings significant complications in the semantics for object-language temporal discourse. The second is that it is unclear whether we can understand the language in which the perdurantist semantic theory is stated.

First problem for perdurance The key perdurantist thesis is that ordinary objects, the subjects of ordinary thought and talk, are spatiotemporally extended worms. So suppose it was true yesterday that Tibbles was purring, but isn't true today. Under what conditions is an atomic predication like 'Tibbles is purring' true at a time? The following won't do:

[31] Clause (i) implies that y only exists in w.

'*a* is *F*' is true at *t* iff the referent of '*a*' is *F*.

If the whole worm that 'Tibbles' refers to is *F*, then this truth-condition makes 'Tibbles if *F*' true yesterday iff it is also true today. This makes it impossible for Tibbles to purr only temporarily. So the perdurantist needs a truth-condition like the following:

'*a* is *F*' is true at *t* iff the *t*-part of the referent of '*a*' is *F*.[32]

This makes it true yesterday that Tibbles was purring iff Tibbles's yesterday-part was purring, and not true today that Tibbles is purring iff Tibbles's today-part is not purring. The properties expressed by predicates like 'is purring' are thus primarily properties of stages of persisting objects.

This truth-condition won't do for all predications. Consider:

Tibbles is a cat.

This should be true whenever Tibbles exists. Applying the truth-condition above gives:

'Tibbles is a cat' is true at *t* iff the *t*-part of the referent of 'Tibbles' is a cat.

But none of Tibbles's *t*-parts is a cat: the ordinary persisting object Tibbles is a cat, not any of his temporal stages. The perdurantist therefore faces a choice. The first option is to deny that Tibbles himself is a cat. On this view, Tibbles is a cat in only the following sense: each of his temporal-parts is a cat. The cost of this view is that the subjects of ordinary thought and talk are not cats, clouds, chairs, humans etc., but fusions of momentary cats, clouds, chairs and humans.

The second option is to adopt a non-uniform account of temporal-modification. On this view, some predications have the truth-condition above, and some have this alternative truth-condition:

'*a* is *F*' is true at *t* iff the referent of '*a*' is *F*.

It's worth noting that this non-uniformity will arise anyway. Consider:

[32] A slightly different truth-condition is the following: the referent of 'Tibbles' is such that its *t*-part is *F*. The difference concerns whether the subjects or predicates of atomic predications vary over time. The following discussion should be insensitive to this difference.

Tibbles has been sleeping for four days.

Tibbles is a temporally extended worm.

Suppose these are both true today (as the second must be, in order for perdurance to be true). Applying the first truth-condition gives:

The today-part of the referent of 'Tibbles' has been sleeping for four days.

The today-part of the referent of 'Tibbles' is a temporally extended worm.

Tibbles's today-part has not been sleeping for four days. And Tibbles's today-part is not temporally extended. So the first truth-condition conflicts with our supposition by making both sentences false. But applying the second truth-condition gives:

The referent of 'Tibbles' has been sleeping for four days.

The referent of 'Tibbles' is a temporally extended worm.

The temporally extended worm Tibbles may well have been sleeping for four days.[33] And that worm is temporally extended. This second truth-condition therefore does not make both sentences trivially false.

We can draw two morals from this. The first is that perdurance complicates our semantic theory. The truth-conditions of some predications turn on the properties of temporal-stages, while the truth-conditions of other predications turn on the properties of worms themselves. The second moral is that this complication is essential to the perdurantist theory; for without it, many of their key theses would be false. If both types of predication were not present in natural language, then the perdurantist would lack the expressive resources to state their view without rendering it trivially false.

Second problem for perdurance It is not clear whether we can understand the perdurantist semantic theory. They offer the r.h.s. of equivalences like the following as an analysis of (the present truth-conditions of) the l.h.s.:

Tibbles is purring iff the now-part of Tibbles is purring.

[33] Presumably, a worm has been sleeping for four days iff each of its temporal-parts over those days is asleep.

Tibbles is sitting iff the now-part of Tibbles is sitting.

For a large class of predications Fa, their perdurantist truth-conditions thus involve applying F not to the referent of a, but to one of its temporal-parts. The difficulty is that although we understand these predicates as they apply to ordinary objects, it is not clear that we understand them as they apply to the temporal-parts of those objects.

What does it mean to say that Tibbles's today-part is purring? The natural answer is: Tibbles is purring. The problem is that this explanation applies the predicate 'is purring' to the persisting four-dimensional object Tibbles, rather than to a stage of that object. It is therefore unclear whether the perdurantist can give an account of the content of their proposed truth-condition for 'Tibbles is purring', other than in terms of the properties of persisting objects. It is therefore unclear whether they can provide an illuminating semantics for truths about persisting objects in terms of truths about their temporal-parts. The suspicion is that when a perdurantist applies an ordinary predicate F to a temporal-part x when stating their semantic theory, the content of this statement is explicable only in terms of the application of F to the persisting object of which x is a temporal-part. I do not claim that this challenge cannot be met, but that it presents a significant obstacle to an informative perdurantist semantic theory.

Perdurance-theory brings two difficulties. Firstly, it complicates our semantic theory. Secondly, it is not obvious that we understand the predicates used in its semantic theory (if the theory is supposed to be informative). We should therefore be reluctant to endorse the predurance-theoretic solution to problems (a)–(d) above. Since we have also rejected counterpart-theory and no alternative solutions are forthcoming, we should be reluctant to endorse the Lewisian Proposal about vague boundaries that generates problems (a)–(d).

3.4 Conclusion

This chapter examined the Lewisian Proposal that vague boundaries result from vagueness about which object's boundaries are in question. We saw two ways of developing this idea in §3.1 and rejected one of them. Before examining the Proposal itself, we defended the Sharpening View's account of vague reference in §3.2. With this account of vague reference in place, we examined three problems for the Lewisian Proposal in §3.3. The first questioned whether it was a genuine solution and exposed its widespread and radical hidden metaphysical assumptions. The second suggested that the Proposal prevents an adequate characterisation of the semantics of self-reference, and hence cannot be extended to cover objects capable of such. The third committed the Lewisian to either counterpart-theory or perdurance. This Lewisian Proposal was thereby revealed as an entrenched component of Lewis's metaphysical framework, not readily separable from the whole. Having already rejected counterpart-theory in §1.1.2.1, we closed by rejecting perdurance. In light of these problems, we should reject the Lewisian Proposal: Unger's and Lewis's puzzles are neither sources nor symptoms of referential vagueness in our names for ordinary objects. We must therefore consider vagueness in mereological, constitutional and loctional concepts themselves, as those concepts apply to ordinary objects. The next chapter defends a view of this kind.

Chapter 4

Identity Conditions and Constitution

The previous chapter rejected Lewisian attempts to resolve the Problem of the Many by postulating referential vagueness in our names for ordinary objects: Unger's puzzle does not induce unclarity about the referent of 'Tibbles', and vagueness about Tibbles's constitution is not a result of such unclarity. This leaves one option: Tibbles's boundaries are vague because mereological and locational vocabulary itself is vague, at least as it applies to ordinary objects. On this kind of view, Tibbles is clearly the unique most cat-like object on his mat. Everything else thereabouts is clearly not a cat, though it's unclear of some things whether they constitute (or compose or make up) a cat. None of Tibbles's Many are cats, or even candidate cats, but merely borderline cases of cat-constituters. Unger's puzzle is one source of this constitutional unclarity (though not uncontroversially of full-blown vagueness). This chapter develops a conception of ordinary objects that accommodates this.

Mark Johnston and E.J. Lowe advocate similar views. §4.1 argues that their views are unsatisfactory because they do not provide a unified solution to the puzzles, but merely an *ad-hoc* collection of theses designed to block the arguments for many cats. §4.2 presents and develops the following thesis about ordinary objects: persistence through change is explanatorily prior to constitution, mereology and location. We argue from this thesis to the following solution to Unger's puzzle:

Tibbles is constituted by each of the best candidates on his mat. §4.3 rebuts several objections to this view. Our solution creates trouble for the idea that Tibbles "inherits" some properties, like mass, from his constituters: surely Tibbles doesn't have several (incompatible) masses. §4.4 develops an account of property-inheritance that avoids this problem. We close with three accounts of vague constitution in §4.5.

4.1 Vagueness in parthood and constitution

Mark Johnston (1992, §4) and E.J. Lowe (1995) defend views of the present kind. Both distinguish constitution from identity and grant that there are many equally good candidates on Tibbles's mat. However, they claim, these candidates are clearly not cats and clearly distinct from Tibbles; they are merely candidates to constitute Tibbles the cat, not candidates to be him. Johnston and Lowe claim that it is vague which candidate constitutes Tibbles, the one and only object on the mat with any claim to *be* a cat.[1]

Lowe claims that this vagueness involves semantic indecision about the extension of 'constitutes', and should be handled supervaluationally. Johnston claims that:

> "The problem of the many simply shows that constitution is a vague relation.... [O]n one legitimate sharpening [Tibbles] is constituted by one of the [candidates], on another, another of the [candidates], and so on. What is important for our purposes is that on no legitimate sharpening is [Tibbles] identical with any one of the [candidates]." (Johnston, 1992, p.100)

Johnston's view is less than perspicuous. Calling constitution a vague relation seems at odds with the linguistic conception of vagueness usually associated with sharpenings and supervaluations.[2] Note however, that Johnston and Lowe agree

[1] Michael Tye (1996) also distinguishes Tibbles from his candidate constituters, though Tye is no supervaluationist.

[2] Williamson (2003) discusses the relationship between supervaluationism and vague properties and relations.

on the following: the candidates are clearly all non-cats, though it is vague which constitutes Tibbles the cat. This section agues that this view is unsatisfactory as it stands.

4.1.1 Why only one cat?

Johnston and Lowe do not provide a unified response to the Problem of the Many, but an *ad-hoc* collection of theses united only by their role in blocking the arguments for many cats. Their view is therefore unsatisfactory.

To see why Johnston and Lowe's view is unsatisfactory, note that the identity/constitution and cat/constituter distinctions alone provide no reason to deny that there are many cats on the mat. They are consistent with:

(1) Each candidate constitutes a cat.

This seems to follow from the following pair:

(2) The candidates are alike in respects relevant to their constituting cats (cat-respects).

(3) Clearly, some candidate constitutes a cat.

Suppose x, y are alike in cat-respects and it's clear that one of them constitutes a cat. Then surely both must constitute cats; for otherwise they would differ in cat-respects because only one would constitute a cat. So (1) seems to follow from (2) and (3).

(1) alone doesn't imply that there are many cats; for the candidates may all constitute the same cat. But since no two candidates spatially coincide, the following imply that no two of them constitute the same cat:

(4) If x constitutes some cat y, then x and y occupy exactly the same place.[3]

(5) No cat occupies more than one region at a time.

We now have a two-step argument from an abundance of candidates to an abundance of cats. The first step concludes that each candidate constitutes a cat. The

[3] We use 'occupies' to mean exact occupation: x occupies region r iff x fills and fits within r.

second step concludes that no two candidates constitute the same cat. The problem for Lowe and Johnston is that they provide no reason to reject either step, or even to endorse the vagueness of constitution. The identity/constitution distinction is consistent with various responses to Unger's and Lewis's puzzles (including Lewis's own). The suspicion is that Lowe and Johnston's solution does not constitute a unified theoretical package. To illustrate, consider the following two responses to our two-step argument for many cats.

First response This response attacks the argument from (2) and (3) to (1): Tibbles's candidates can be alike in cat-respects despite none clearly constituting a cat. All that's required in order for it to be clear that some candidate constitutes a cat, Lowe and Johnston may claim, is that each candidate borderline constitutes a cat; no candidate clearly does so, though none clearly fails to do so either. A supervaluationist logic on which $\Delta\exists xA$ doesn't imply $\exists x\Delta A$ is obviously congenial to this. (2) and (3) imply that each candidate clearly constitutes a cat only if 'constitutes' is not vague. The vagueness of constitution therefore invalidates the first step of our argument for many cats.

There are three problems for this response. First, the identity/constitution distinction neither implies nor suggests that 'constitutes' is vague. It therefore provides no reason to believe that the argument is invalid. Second, even granting that constitution is vague, Lowe and Johnston provide no reason to think that (1) is false. Our argument shows that either constitution is vague or each candidate constitutes a cat. No basis to prefer one disjunct to the other has been supplied. Third, this disjunction is *inclusive*: it is consistent with the vagueness of constitution and the identity/constitution distinction that many vaguely constituted cats are on the mat. Lowe and Johnston provide no reason to think otherwise.[4]

Second response This response attacks the second step of our argument for many cats. Since (4) and (5) seem too firmly embedded in our conception of the material world to be plausibly denied, this response doesn't reject, but modifies them. It

[4] Matti Eklund (2008) raises a similar worry about appeal to ontological vagueness in response to the Problem of the Many.

does so by denying that cats occupy space in the same way as their constituters: cats occupy space *by* being constituted by objects that occupy space. The spatial properties of cats are thus relational properties: cat x occupies region r relative to its constituter y; this relation obtains iff y both constitutes x and occupies r in the primary, non-relativised, sense.

On this view, (4) and (5) become:

(4′) If x constitutes some cat y, then: y occupies r relative to x iff x occupies r.

(5′) No cat occupies, relative to anything that constitutes it, more than one region at a time.

(4′) is trivial, given the proposed account of spatial occupation by cats. (5′) follows from the (plausible) assumption that no cat-constituter occupies more than one region at a time. From the supposition that each candidate constitutes the same cat—i.e. that each candidate constitutes Tibbles—(4′) implies:

> For some distinct regions r, r' and distinct candidates c, c', Tibbles occupies r relative to c, and r' relative to c'.

Since this is consistent with (5′), it doesn't follow that the candidates don't all constitute the same cat.

This second response suffers the same problem as the first. The identity/constitution distinction provides no reason to endorse it. That distinction doesn't imply that cats and their constituters occupy space in the fundamentally different ways that this response claims they do.

In sum, then, although the identity/constitution distinction does create logical space for (at least) two responses to the Problem of the Many, it provides no reason to endorse either of them. It doesn't even provide reason to deny that there are many cats on Tibbles's mat. Lowe and Johnston therefore do not provide unified theoretical packages. This chapter seeks to do better. Our goal is a conception of ordinary objects from which a solution to Unger's and Lewis's puzzles emerges naturally, alongside the identity/constitution distinction, unclarity in constitution and Tibbles's uniqueness.

4.2 The Proposal

This section develops a conception of ordinary objects and applies it to Unger's puzzle of too many best candidates. §4.2.1 outlines the core idea. Central to this idea are criteria of identity. §4.2.2 presents two kinds of identity criterion. §4.2.3 develops our proposal in accord with the first kind of criterion and applies it to Unger's puzzle. §4.2.4 does the same with the second kind of criterion. §4.2.5 then argues that we needn't choose between these proposals, and §4.2.6 closes by examining the concepts of matter and constitution employed throughout our discussion. The next section turns to some objections.

4.2.1 Objects and change

Our primary interest is in ordinary objects, the kind of object singled out in ordinary thought and talk, and the inhabitants of the ordinary macroscopic world. We're also interested in the ordinary parts of these objects, e.g. the hearts and lungs of animals, the cells of plants, and the legs of tables, because the Problem of the Many arises for them too. So what is it to be an ordinary object?

Our proposal develops a broadly Aristotelian answer to this question: an ordinary object is a subject of change; different kinds of object are subjects of different kinds of change. This idea is captured by associating each ordinary kind K with an identity condition that determines what changes Ks survive; identity conditions determine the histories and futures of ordinary objects.

What is the content of our claim that to be an ordinary object is to be a subject of change? A strong form of this thesis takes it as a (conceptual? metaphysical?) analysis of belonging to the category of ordinary objects. However, the following weaker thesis will suffice for our purposes whilst allowing us to avoid difficult questions about just what such analyses amount to:

> Identity conditions, history and change are explanatorily prior to the constitution, mereology, and spatial location of ordinary objects.

The idea is that every fact about the constitution, mereology and location of ordinary objects is explicable in terms of the kinds of changes that those objects do and don't survive (in combination with other contingent facts). An object's identity condition determines its path through time and space; it's constitution, mereology and most other properties follow from what happens along this path.

David Wiggins defends a view along similar lines:

> "Suppose I ask: Is *a*, the man sitting on the left at the back of the restaurant, the same person as *b*, the boy who won the drawing prize at the school I was still a pupil at early in the year 1951? To answer this sort of question is surprisingly straightforward in practice.... Roughly, what organizes our actual method is the idea of a particular kind of continuous path through space and time the man would have had to have followed in order to end up here in the restaurant.... Once we have dispelled any doubt whether there is a path in space and time along which that schoolboy might have been traced and we have concluded that the human being who was that schoolboy coincides with the person/human being at the back of the restaurant, this identity is settled.... The continuity or coincidence in question here is that which is brought into consideration by what it is to be a human being.... The contention is... that to determine correctly the answer to our continuity question, the question about the traceability of things through their life-histories, precisely *is* to settle it that, no matter what property ϕ is, *a* has ϕ if and only if *b* has ϕ." (Wiggins, 2001, pp.56–7)

I include the last sentence because one property of *b* is the property of *being b*; thus to settle how to trace the histories of *a* and *b* is, in part, to settle whether *a* is *b*. And how to trace the histories of *a* and *b* is settled by what kind of objects they are, human beings in Wiggins's example.

How does this help with the Problem of the Many? There's more detail in §§4.2.3–4.2.4, but a preview may be helpful. By determining what kinds of change it survives, an object's identity condition associates it with a path through space and time. The Problem of the Many shows that a single object's path may exhibit a branching structure, passing through several nearly coincident spatial regions at a single time, when examined on a sufficiently small scale. When this happens, the ordinary object with this branching history is simultaneously constituted by the occupants of each of these regions; these occupants are its Many. Our thesis about the relative explanatory priorities of change and constitution even allows us to argue for this last claim.

Let T be a cat on Tibbles's mat at time t_1. Let l_2 be one of the candidates on the mat at the later time t_2. The question is: does l_2 constitute T at t_2? The answer will be positive if the change c that T would have to survive in order to come to be constituted by l_2 at t_2 is a kind of change that cats do survive. We may safely assume that T does survive some change c^*, as a result of which it comes to be constituted by some other candidate l_2^* at t_2. Given how similar the candidates on the mat at any given time are to one another—Unger's puzzle only arises if l_2 and l_2^* are alike in cat-respects—the changes c and c^* will also be very similar: c and c^* are alike in respects relevant to whether they are the kind of change survived by cats. But then, how could T survive c^* but not survive c? If either is the kind of change that cats survive, then surely they both are. Since, we assumed, T does survive c^*, it also survives c. So T survives both changes, as a result of which it comes to be constituted by both l_2 and l_2^* at t_2. Generalising: all the candidates on the mat at t_2 simultaneously constitute one and the same cat. A similar argument gives the same result for t_1. Hence T is, at every time, constituted by every candidate then on the mat: T is Tibbles, the one and only cat ever on the mat.

The moral is that if objects are individuated by their histories rather than by their microscopic constituents, then many equally good candidate cat-constituters needn't correspond to many cats. They may instead all constitute a single cat. The Problem of the Many is a symptom of an overemphasis on mereology in the ontology of ordinary objects.

To make good on these claims, more detail is required about identity criteria and constitution. The next section introduces two kinds of identity criterion. Subsequent sections use them to develop two versions of our proposal.

4.2.2 Two kinds of identity criterion

Three putative examples of identity criteria are prominent in the literature; those for sets, directions and cardinal numbers:[5]

(Extensionality) $(\forall x : \text{Set}(x))(\forall y : \text{Set}(y))(x = y \leftrightarrow \forall z(z \in x \leftrightarrow z \in y))$

(Dir) $\forall x \forall y(d(x) = d(y) \leftrightarrow x \| y)$

(HP) $\forall F \forall G(\#(F) = \#(G) \leftrightarrow F1\text{–}1G)$

'Set' is a predicate true of exactly the sets; '*d*' denotes the function from lines onto their directions, and '$\|$' the relation of parallelism; '#' denotes the function from concepts or properties onto their cardinalities, and '1–1' the relation of one-one correspondence. Thus (Extensionality) says that sets are identical iff they have exactly the same members; (Dir) says that lines have the same direction iff those lines are parallel; and (HP) says that concepts have the same cardinality iff those concepts are in one-one correspondence.

Abstracting away from these specific examples (and ignoring the higher-order quantifiers in (HP)), there are two kinds of statement here:

One-level $(\forall x : F(x))(\forall y : F(y))(x = y \leftrightarrow R(x, y))$.

Two-level $\forall x \forall y(f(x) = f(y) \leftrightarrow R(x, y))$.

The labels are from Williamson (1990, §9.1). In a true one-level criterion, R is an identity condition for Fs. In a true two-level criterion, R is an identity condition for $f(x)$s.[6] The formal properties of identity require that identity conditions are equivalence relations.

One-level and two-level criteria obviously have different logical forms. But for our purposes, the question of logical form is secondary to another difference. In a one-level criterion, R is a relation on the very objects whose identity is at issue on the left. In a two-level criterion, by contrast, R need not be. Witness that coextensiveness is a relation on sets, though parallelism and one-one correspondence

[5] $(\forall x : A)$ is a quantifier restricted to satisfiers of A.

[6] We use 'identity criterion' for statements of either of the forms above, and 'identity condition' for the relations on the r.h.s. of true identity criteria. We'll also call (putative) identity conditions for Fs one-level or two-level, according as to whether they are relations on Fs or not.

are relations on lines and concepts respectively, not on directions and cardinalities. Hence the choice between one- and two-level criteria amounts to the following: is the identity condition for Ks a relation on Ks?

Actually, this isn't quite right. An identity condition R for Ks in a two-level criterion *may* be a relation on Ks. The key difference between the two kinds of statement is that the logical form of a two-level criterion for Ks does not require that R be a relation on Ks, whereas the logical form of a one-level criterion does. For simplicity, we'll confine ourselves to genuinely two-level criteria in the sequel: two-level criteria in which the putative identity condition R is not a relation on $f(x)$s.

We want to use identity criteria to explicate the idea that an ordinary object is a subject of change: the identity condition R for Ks is explanatorily prior to the constitution, mereology and locations of Ks. The choice between one-level and two-level criteria thus amounts to a choice concerning which kinds of change to regard as prior to constitution, which kinds of change explain the the facts about constitution. A one-level version of our proposal concerns changes in ordinary objects themselves. A two-level version concerns changes in something else, whose properties and relations are systematically correlated with identity amongst ordinary objects.

We'll ultimately prefer a one-level view, though we won't reject two-level criteria outright. Note however that either kind of criterion will suffice. Our proposed solution to the Problem of the Many is insensitive to this nuance of formulation. The explanatory priority of change and history over constitution and mereology is what's doing the work. The following sections exhibit this by developing both one-level and two-level versions of our proposal and applying them to Unger's puzzle.

4.2.3 The one-level proposal

The one-level variant of our proposal claims that, for each ordinary kind K, some unique equivalence relation R_K on Ks satisfies:

(L1-K) $$(\forall x : Kx)(\forall y : Ky)(x = y \leftrightarrow R_K(x, y))$$

Applied to cats, we get:

(L1-Cat) $(\forall x: \text{Cat}(x))(\forall y: \text{Cat}(y))(x = y \leftrightarrow R_c(x,y))$

Cat x is identical to cat y iff x bears R_c to y. R_c thus determines what changes cats survive. What kinds of change are these? Two issues arise. Firstly, what are the subjects of these changes? Secondly, what kinds of change are these? We address these in turn.

Since (L1-Cat) is a one-level criterion, R_c is a relation on cats. It holds between a cat x and a cat y iff the changes that x would have to survive in order to be y are of the kind that cats do survive. The one-level proposal thus prioritises changes in cats themselves—as opposed to changes in their underlying matter or the regions they occupy—over the constitution and mereology of cats. R_c bears on the survival of cats through changes in their matter only insofar as those material changes correlate with changes in cat(s). This addresses the first question at the end of the previous paragraph. So let's turn to the second: what kinds of change do cats survive?

Set to one side the question of whether a complete and informative account of R_c is possible. We'll return to that in §4.3.5. Our grasp on R_c is firm enough for present purposes, regardless of whether English contains a (possibly complex) expression coextensive with it. Cats persist through hair loss, purring, falling asleep, pouncing, digesting and countless other ordinary and familiar kinds of change. They don't survive drowning, starvation, squashing and the like. Note that these are all primarily macroscopic changes; changes whose subjects are *cats*.

Unger's puzzle arises because our concept 'cat on the mat' fails to determine a unique microscopically individuated portion of matter as the constituter of the ordinary macroscopic object in question. Likewise, our concept 'the loss of hair h' fails to determine a unique microscopically individuated occurrence as the material basis for the ordinary macroscopic change in question. Our one-level proposal puts this to work in a solution to Unger's puzzle.

4.2.3.1 Applying the one-level proposal

This section applies the one-level proposal to Unger's puzzle of too many candidates. For simplicity, we'll assume that cats are constituted by lumps of matter and restrict attention to those lumps on Tibbles's mat that are candidate cat-constituters. These notions of matter and constitution are elaborated in §4.2.6.

Let T be a candidate lump of matter on the mat at t_1. Suppose for simplicity that it's the only candidate then on the mat, and hence that it then constitutes Tibbles. Let T_2, T_2^* be two candidates lumps on the mat at the later t_2. Suppose for simplicity that they're the only candidates then on the mat. Suppose Tibbles clearly survives from t_1 to t_2: the changes that occur on his mat over that duration aren't the kind that destroy cats. We want to show that Tibbles is constituted by both T_2 and T_2^* at t_2.

Since Tibbles survives from t_1 to t_2, we know that he survive some change c, as a result of which he comes to be constituted by at least one of T_2, T_2^* at t_2. Suppose that it's T_2. Now consider the change c^* that Tibbles would have to survive in order to come to be constituted by T_2^* at t_2. How could Tibbles survive c but not c^*? Since T_2 and T_2^* are very similar, c and c^* are too. If it's implausible that only one of T_2 and T_2^* constitutes a cat, then surely it's also implausible that only one of c and c^* is the kind of change that cats survive. The ordinary changes and process that cats survive though—the loss of a hair, purring, walking, pouncing and so on—don't seem to distinguish between objects as similar as T_2 and T_2^* in respect of which comes to constitute the cat that participates in them as a result of its undergoing that change. Both c and c^* are equally good candidates to be changes that Tibbles survives. Since he clearly survives one of them, surely he must survive both. Hence T_2 and T_2^* both constitute Tibbles at t_2.

Nothing in this argument turned on the assumption that T is the only candidate-constituter on the mat at t_1. So we can generalise: for any candidate x on the mat at any time t, and for any candidate y on the mat at any other time t', x constitutes the same cat at t as y does at t'.

This argument shows that any pair of candidates drawn from different times constitute the same cat as one another. It doesn't follow that there's only one cat on the mat; for all we've said so far, each of the candidates may constitute more than one cat. T, for example, may constitute two cats at t_1, one of which is constituted by T_2 and t_2, and one of which is constituted by T_2^* at t_2. Similarly, even if T_2 and T_2^* both constitute the same cat, they may both constitute more than one. We'll now argue that this isn't the case.

Our key thesis is that change is explanatorily prior to constitution. One aspect of this is that all the facts about constitution are explicable in terms of facts about change. So if any candidate constitutes more than one cat, then that's explicable in terms of the kinds of change that cats survive. If there's no such explanation, then no candidate constitutes many cats; in which case there's only one cat on Tibbles's mat. What would such an explanation be?

An explanation in terms of change for there being more than one cat on the mat would involve a pair of changes such that, from the occurrence of both on the mat, it follows that there's more than one cat; a pair of changes that couldn't both be survived by a single cat. Ordinary changes like the loss of a hair don't seem to provide this. The only candidates I can find are pairs like:

> The change a cat undergoes when it ceases to be constituted by T and comes to be constituted by T_2 and not by T_2^*.
>
> The change a cat undergoes when it ceases to be constituted by T and comes to be constituted by T_2^* and not by T_2.

The supposition that a cat survives both these changes is inconsistent. So no cat can do so. So if cats do survive these changes (and both occur on the mat), then there were two cats on the mat at t_1, both constituted by T. This kind of suggestion faces four problems.[7]

Firstly, it's not obvious that these are genuinely one-level changes because they're only specifiable using features of matter. Secondly, since both changes are specifiable only using constitutional vocabulary, they conflict with our thesis of the ex-

[7] This objection doesn't require the claim that cats do survive these changes, or that they do occur on the mat. All that's required is that we can't rule out that they do, and hence can't rule out there being many cats.

planatory priority of change over constitution.[8] Thirdly, these changes are so similar that it's unclear how any cat could survive only one of them. But since no cat can survive both changes, it follows that no cat survives either; neither is the kind of change that cats survive. Fourthly, these changes appear too artificial and gerrymandered to be plausibly taken as characteristic of cats.

Given these difficulties, there seems no explanation in terms of change for how any candidate could constitute more than one cat. So no candidate does. Tibbles is therefore the only cat ever on the mat. At each time, he's constituted by each candidate then on the mat.

This one-level proposal employs the phenomenon that generates Unger's puzzle as part of a solution. That puzzle arises because our ordinary sortal concepts don't make sufficiently fine-grained distinction amongst lumps of matter in respect of which lumps constitute the satisfiers of those concepts. Likewise, our sortal concepts don't distinguish between the changes that cats would have to survive in order to come to be constituted by some one later candidate in preference to any other. Since those changes have equally good claim to be survived by Tibbles and he survives at least one of them, Unger's reasoning concludes that he survives them all. The result is that Tibbles comes to be multiply constituted by all the best candidates on his mat.

4.2.4 The two-level proposal

Recall the general form of a two-level criterion:

$$\forall x \forall y (f(x) = f(y) \leftrightarrow R(x, y))$$

Two-level criteria treat the $f(x)$s as invariants across R-connected series of xs. Since our interest is in ordinary objects, the $f(x)$s will be ordinary objects. Thus ordinary objects are invariants across R-connected series of some other kind of entity. What kind of entity?

[8] One might respond to these first two problems by claiming that although the changes in question are only specifiable in terms of constitution and features of matter, this is simply a function of the expressive limitations of natural language, rather than a reflection of any deep metaphysical fact. But the onus is on our opponent to justify this perspective. Why should we believe that appearances are misleading in the way that they claim?

The quantifiers in our two-level criteria range over a kind of entity such that, no matter how unlike one another members of that kind may be, they "support" the same ordinary object iff they stand in R. Changes in such entities that don't bear on R, don't bear on their relations to ordinary objects. Two questions arise. Firstly, what are these entities? And secondly, what is this "support" relation?

Our concern is with the constitution of cats by the lumps on Tibbles's mat. It's therefore natural to take the domain of f as lumps of matter, its range as cats, and the relation between x and $f(x)$ as the constitution relation. On this view, an ordinary object is an invariant across changes in various lumps of matter, its constituters over time. Our proposal then becomes the thesis that variation in the properties of matter is explanatorily prior to the constitution and mereology of ordinary objects; the r.h.s. of a two-level criterion is explanatorily prior to the left.

This won't quite do because ordinary objects are typically constituted by different matter at different times. So we'll take the quantifiers in our two-level criteria to range over ordered pairs $\langle l, t \rangle$ of lumps l and times t, and take f as the (partial) function from $\langle l, t \rangle$ pairs onto the object (if any) constituted by l at t. Then we can simulate cross-time variation in the relata of R without having to complicate the criteria by introducing extra argument places for times on the right. We will however frequently write as if R held amongst different lumps at (and across) times, and as if f_c mapped different lumps onto different cats at different times. The difference is merely notational.

Two problems remain. Firstly, different kinds of object survive different kinds of change in their matter. Secondly a single lump may constitute several objects at once (e.g. some wool may constitute a thread and a jumper). The following proposal avoids both problems.

Let K be any ordinary kind. Let f_K be the (partial) function from $\langle l, t \rangle$ pairs onto the K (if any) that l constitutes at t. Then some unique relation R_K satisfies:

(L2-K) $$\forall x \forall y (f_K(x) = f_K(y) \leftrightarrow R_K(x, y))$$

Applied to cats:

(L2-Cat) $$\forall x \forall y (f_c(x) = f_c(y) \leftrightarrow R_c(x, y))$$

R_c determines a range of paths through various lumps of matter over time. Any lumps that lie on these paths constitute the same cat (whilst they lie on that path). R_c thereby determines the survival of cats through changes in their matter. Before applying this to Unger's puzzle, following are six comments by way of elucidation.

First comment Although (L2-K) makes it formally possible to identify Ks with equivalence classes of R_K-related $\langle l, t \rangle$ pairs, it does not mandate doing so. Neither should we wish to. Cats are ordinary concrete objects, not set-theoretic constructs. They do not have members, they are not abstract, and (according to our proposal) they are not "formed" from or dependent upon lumps of matter, times or $\langle l, t \rangle$ pairs. An equivalence class of R_K-related $\langle l, t \rangle$ pairs is, at best, a formal model of a cat.

Second comment The identity/constitution distinction follows directly from focusing on genuinely two-level criteria. It also follows from the fact that f_K maps distinct $\langle l, t \rangle$ pairs onto the same K: since cats are constituted by different lumps at different times, constitution is not identity; cats are not identical to their matter.

Third comment Although f_K was introduced as the function from $\langle l, t \rangle$ pairs onto the K constituted by l at t, this is not a definition of f_K. We want to leave it open how closely f_K coincides with our ordinary notion of constitution. (§4.2.6 discusses this notion.) In particular, we want to leave it open whether a single ordinary object might be simultaneously constituted by several lumps of matter. In part, that's because our solution to Unger's puzzle requires that they can be. But it also coheres with our thesis of the explanatory priority of change and persistence over constitution.

Fourth comment The fundamental difference between (L1-Cat) and (L2-Cat) concerns whether R_c, the identity condition for cats, is a relation on cats or on lumps of matter. (L2-Cat), unlike (L1-Cat), treats cats as invariant through changes in non-cats, specifically in their matter. On this view, the changes undergone by cats bear on their persistence only insofar as they correlate with changes in their matter.

One benefit is that the two-level proposal avoids the appearance of bootstrapping that may attend the one-level solution.

Fifth comment Peter Simons (2000, 2008) also endorses two-level identity criteria for ordinary objects (and continuants more generally). His view differs from ours in three respects. (i) Simons takes the identity condition for *K*s as a relation on events, not on matter or other continuants. (ii) Simons takes the notion of an invariant with more metaphysical seriousness than we do. His two-level criteria reflect a view of continuants as abstractions from, and dependent upon, metaphysically fundamental events. Our two-level proposal involves no theses about the non-fundamentality of ordinary objects, or their dependence on the domain of f_c. We claim only that the relations on that domain are what determine the survival of cats. (iii) Simons endorses a different account of property-possession to our own, but we won't go into that here.

Sixth comment E.J. Lowe writes:

> "[T]he parallelism of lines can provide a criterion of identity for the *directions* of lines only because directions are ontologically (and indeed conceptually) dependent on lines in a way that lines are not on directions. But this immediately raises a difficulty for anyone seeking to extend [two-level criteria] to names of what we might, in an Aristotelian vein, call (primary) *substances*, since these (assuming they exist) are precisely the objects standing in no such relationship of ontological dependency to other objects." (Lowe, 1989, p.4; original italics)

If there are any substances, cats are presumably amongst them. And our conception of ordinary objects is closely akin to the Aristotelian notion of substance. So Lowe's objection had better be unsound. It is.

Grant for the sake of argument (the highly dubious assumption) that there is a contentful notion of a specifically ontological kind of dependence. Our two-level proposal involves no explicit theses about such dependence. It merely posits (i) a structural thesis connecting cat-constitution and a relation on matter, and (ii) the priority of identity conditions over constitution and mereology. Lowe's objection therefore carries weight only given an argument from our proposal to the relevant claims about dependence. Lowe offers no explicit argument, but the following is suggestive:

> "[A]n acceptable criterion of identity for ϕs... should reveal, in an informative way, what ϕ-identity 'consists in' (to use Locke's well-worn phrase). In particular, then, it should reveal to us wherein ϕ-identity differs from ψ-identity." (Lowe, 1991, p.193)

Suppose that if K-identity consists in a relation on Fs, then Ks are ontologically dependent on Fs. It follows that if cats are ontologically independent, then cat-identity doesn't consist in a relation on non-cats. So if Lowe's claim about identity criteria is correct, then (L2-Cat) is not an acceptable identity criterion for cats.

If sound, the argument in the previous paragraph shows that (L2-Cat) does not provide an analysis of cat-identity, or an account of what cat-identity consists in. This doesn't undermine the two-level view for two reasons. Firstly, we could regard the r.h.s. of (L2-Cat) as an analysis of its l.h.s. without regarding R_c as an analysis of cat-identity; for in the l.h.s. '$f_c(x) = f_c(y)$', the identity sign is flanked by complex terms formed using functional signs that denote a constitution relation, and variables ranging over matter. If the two-level theorist presents R_c as an analysis of anything, it's as an analysis of this complex statement, not bare identity. Secondly, the two-level theorist is not compelled to regard the r.h.s. of (L2-Cat) as an analysis of anything. That criterion may instead be taken to contribute to a better understanding of the kind *cat*, identity, constitution and material change by exposing their inter-connections: (L2-Cat) imposes a structure on these connections, which information about R_c imbues with content. Lowe employs too strong a notion of ontological dependence if the *in*dependence of ordinary substances is incompatible with systematic correlations between their persistence and the properties of matter. Surely the existence of such connections is uncontroversial. Lowe's objection therefore fails.

Now the two-level proposal is in place, let's apply it to Unger's puzzle.

4.2.4.1 Applying the two-level proposal

We want to use (L2-Cat) to show that there's never more than one cat on Tibbles's mat. This section presents two strategies. The first requires slightly more substantive metaphysical assumptions than the second.

First strategy We begin by arguing that if there's ever n cats on the mat, then there's never more than n.

Suppose there are n cats on Tibbles's mat at t_1: $\langle l_1, t_1 \rangle, \ldots \langle l_n, t_1 \rangle$ each bears R_c to something, though not to one another. Suppose also that there are $n+1$ cats on the mat at some other time t_2: $\langle l_1^*, t_2 \rangle, \ldots \langle l_{n+1}^*, t_2 \rangle$ each bears R_c to something, though not to one another. Since there are more l^*s than ls, some pair $\langle l_i^*, t_2 \rangle$ either (i) bears R_c to the same $\langle l_j, t_1 \rangle$ pair as some other $\langle l_k^*, t_2 \rangle$ pair, or (ii) bears R_c to no $\langle l_j, t_1 \rangle$ pair. But (i) implies that l_i^* and l_k^* constitute the same cat at t_2; in which case there aren't $n+1$ cats on the mat at t_2, contrary to our second supposition. And (ii) implies that l_i^* constitutes a cat that wasn't on the mat at t_1, which, we may assume, is obviously false: the occurrences on Tibbles's mat didn't bring a cat into being or fuse Tibbles with some cat not previously on his mat. So our two suppositions are incompatible. So if there are ever n cats on the mat, then there are never more than n cats on the mat.

So far, so good. But we want to show that there's never more than one cat on Tibbles's mat. This follows from:

Independence Whether R_c holds between $\langle l, t \rangle$ and $\langle l^*, t^* \rangle$ is independent of what happens outside of the duration from t to t^*. Whether l constitutes the same cat at t as l_* does at t^* turns only on what happens between t and t^*.

Possible Uniqueness Either (i) events could possibly unfold so that there's only one candidate on Tibbles's mat at some point in the future, or (ii) there's a possibility w in which (a) there's only one candidate on Tibbles's mat and (b) events could possibly unfold from w to give the actual present situation.

To understand Possible Uniqueness, think of reality as a branching structure. Different branches represent different possible complete histories (including futures).

Shared portions of branches represent shared portions of possible histories. Possible Uniqueness says that each branch intersects with one where there's only one candidate on Tibbles's mat: each history shares a portion with one where there's only one cat on Tibbles's mat; whatever the situation on Tibbles's mat, it could either lead to a future situation or be the result of a (merely possible) past situation in which there's only candidate on the mat.[9] Unger's argument suggests that such one-candidate situations are unlikely to be actual, not they are impossible. The argument above shows that any branch with only one candidate at some point has only one cat at all points. Independence then extends this to any connecting branch. Possible Uniqueness says that every branch connects with some such branch. So there's never more than one cat on Tibbles's mat.

Although Possible Uniqueness and Independence seem natural, they are non-trivial and I know of no direct arguments for them. In support of Possible Uniqueness, it's hard to imagine instances of Unger's puzzle that couldn't be reached from, or couldn't unfold into, some situation in which there's only one candidate: what would Tibbles have to now be like in order to prevent there ever being a time at which there's only one candidate? And in support of Independence, we might claim that R_c marks an intrinsic similarity between its relata. For those suspicious of these theses however, a slightly different strategy is available.

Second strategy The Problem of the Many arises only when the candidates are alike in respects relevant to cat-constitution. In a two-level setting, this amounts to the candidates being alike in respects relevant to R_c. Since R_c is an equivalence relation, each candidate bears R_c to itself. So if candidate x doesn't bear R_c to candidate y, then they differ in at least one respect relevant to cat-constitution: only y bears R_c to y. Since they don't differ in such respects, each candidate bears R_c to each other. Similarly, if x but not y bears R_c to some future-candidate z, then x and y differ in a respect relevant to cat-constitution, namely, the bearing of R_c

[9] A weaker claim that would suffice is: either (i) each history shares a portion with one in which there's only one cat; or (ii) each history shares a portion with a history that shares a portion with one in which there's only one cat; or (iii) each history shares a portion with a history that shares a portion with a history that shares a portion with one in which there's only one cat; or (iv) each history shares....

to z. Since the candidates are alike in those respects, any two present-candidates bear R_c to the same future-candidates. Since R_c is an equivalence relation, they also bear it to one another. So they constitute the same cat, and there's never more than one cat on the mat. At every time, the cat on the mat is constituted by each best candidate-constituter on the mat.

Two features combine here. The first is that the candidates are alike in cat-respects. The second is that we've fleshed out cat-respects using a relation. This second feature forces us to consider a range of candidates—those to which a given candidate bears R_c—when assessing resemblance in cat-respects. The first then ensures that the candidates bear R_c to the same ranges of candidates. The similarities between candidates that give rise to Unger's puzzle thereby contribute to resolving it.

A slightly different way to see this last point is as follows. Suppose that there are two earlier candidates T_1, T_2 and two later candidates T_3, T_4. If R_c holds from, say, T_1 to T_3 but not to T_4, then this provides an example of a change in T_1's matter that cats don't survive. But the similarity of T_3 to T_4 makes this implausible if that similarity is close enough to undermine the claim that only one of them constitutes a cat. So T_1 and T_4 constitute the same cat iff T_1 and T_3 do. Since T_1 constitutes the same cat as either T_3 or T_4, it must constitute the same cat as them both. Likewise with T_2 in place of T_1. So T_1 and T_2 stand in R_c, and therefore constitute the same cat. Generalising: there's never more than one cat on the mat, and it's constituted at each time by each candidate then on the mat.

4.2.5 How many levels?

We've now got two responses to Unger's puzzle in place. Isn't this one too many? Which should we choose? Although §4.3.5 presents some (inconclusive) reasons to prefer the one-level proposal, the two views aren't competitors.

Our core thesis is the explanatory priority of change over constitution. The one-level and two-level views differ over which kinds of change they claim take priority. But the one-level view needn't claim that changes in cats take priority over changes in their matter. And the two-level view needn't claim that changes in matter take priority over changes in cats. We might instead see both kinds of change as explanatorily prior to constitution. This requires that one-level and two-level identity criteria don't deliver conflicting results about the persistence and constitution of cats. But since there's no reason to think that they will, we're not forced to decide between these two views.

4.2.6 Matter and constitution

We've made free use of the idea that ordinary objects are constituted by lumps of matter. Can we be sure that lumps of matter exist? And what is this constitution relation? This section addresses these questions in turn.

We needn't assume the existence of lumps of matter in any significant sense. Maybe fundamental reality consists of pluralities or aggregates of microscopic particles, regions of spacetime, or something else entirely. Our use of 'matter' is best understood as a placeholder for whatever physical stuff ordinary objects are ultimately made out of. And if there's no absolutely fundamental stuff, then an appropriate choice of level will suffice. Or we could simply re-parse our discussion in terms of the occupation of spatiotemporal regions by objects.

Our use of 'constitution' also comes with minimal theoretical baggage. It's best seen as a technical term for an ordinary concept. It's beyond doubt that ordinary objects are in some sense made from other (smaller) entities. We use 'constitution' to denote whatever relation occupies this "making up" role, without adopting any substantive metaphysical views about it, other than that several entities can bear this relation to a single ordinary object (though see §4.5.1). Our key thesis is that, whatever relation constitution is and whatever ordinary objects are constituted by, persistence and change are explanatorily prior to constitution and mereology.

4.2.7 Identity criteria: concluding remarks

This section presented two solutions to Unger's puzzle of many best candidates. Although their details differ, the core idea is the same. Each ordinary kind K is associated with an identity condition that determines what changes its members survive. Unger's puzzle arises when these changes don't determine a unique lump of matter as the best candidate to constitute a given *K*. Since these lumps are all on a par and at least one constitutes a *K*, they all do. But because they're so similar—they've all been selected as *K*-constituters by the identity condition of an individual *K*—these lumps all simultaneously constitute one and the same *K*.

We can now diagnose the fundamental flaw in the argument from many lumps to many cats: it assumes that objects are built up from or individuated by their microscopic constituents in an ontologically significant sense. On this kind of view, to be an ordinary object (of a kind *K*) is to be made out of smaller objects in an appropriate way. This makes it inevitable that differences in suitably arranged small objects correlate with differences about which ordinary objects they make up. Unger's puzzle arises when many such suitably arranged collections almost, but not quite, coincide.

Jettisoning this conception of objects invalidates the argument from many candidates to many cats. We've also seen that it brings positive arguments against there being many cats. An overemphasis on mereology in contemporary ontology disguises this fundamental mistake.

Classical extensional mereology, and variants thereof, provide prominent examples of this flawed conception of the relationship between ordinary objects and their microscopic parts.[10] But the mistake is not confined to these views. Fine (1999, 2008) proposes a hylomorphic view according to which an object's fundamental nature—its real definition, or essence—is given by a list of its parts and the form, or universal, they instantiate To be a cat, for example, is to be a collection of atoms in the form of a cat. Although Fine's approach is radically unlike classical mereology, it shares the same flaw. If an ordinary object's underlying nature is given by a list of its microscopic constituents, then different lists must correspond to different objects. The inclusion of a form as an additional item on the list makes no difference to this. The lesson of Unger's puzzle is that any conception of objects, including Fine's hylomporhism, that associates each ordinary object with a unique collection of microscopic constituents is false. To be an ordinary object is not to be made out of appropriately arranged stuff, but to survive through certain sorts of change.

[10] See Simons (1987) for discussion of classical mereology.

4.3 Objections

This section responds to seven objections to our proposal. §4.3.1 considers two objections arising from the principle of Unique Constitution. §4.3.2 addresses an objection we raised against a similar proposal in §1.1.4.2. We discuss a purported similarity between the Problem of the Many and fission and fusion in §4.3.3, and Lewis's scepticism about the cat/cat-constituter distinction in §4.3.4. §4.3.5 asks whether an informative statement of either kind of identity criterion is possible. §4.3.6 closes by examining the claim that identity criteria attempt the impossible, namely an analysis of identity.

4.3.1 Unique Constitution

Consider:

Unique Constitution (UC) No cat can ever be constituted by more than one lump of matter at a time.

Our solution is incompatible with UC. This section considers two problems this creates.

4.3.1.1 First objection

The first objection runs as follows: since UC is true and incompatible with our proposal, that proposal is false. We should not find this compelling. Simply asserting a theoretical claim like UC without supporting argument carries no suasive force. Following are two arguments for UC (and responses).

The first argument for UC appeals to ordinary usage of constitutional vocabulary. The idea is that UC is a deeply entrenched part of ordinary discourse, and therefore carries strong intuitive support. Evidence comes from our use of definite descriptions, e.g.: 'the clay that used to make up the statue is now a set of dishes'. Unless the statue was constituted by a unique lump of clay, the initial description in this example is improper and the whole sentence therefore untrue. But, we may assume, surely it was true; such claims are part of what fix the meaning of 'constitutes'. Our response to this argument must await §4.5.1. That section shows that our proposal doesn't make definite descriptions like 'the matter of Tibbles' improper and can therefore accommodate this kind of linguistic consideration.

The second argument for UC comes from §4.1. Cats are located wherever their constituters are located. Cats are also only located in one place at a time. But only one lump can occupy a place at a time.[11] So cats cannot be simultaneously constituted by multiple lumps.[12]

There are two responses to this argument. One must await §4.4. There we provide an account of Tibbles's location and other inherited properties that is compatible with our proposal. The other response can be stated now: our proposal was not simply that change is prior to constitution, but that it is also prior to location. Exact parallels of our arguments for multiple constitution show that Tibbles multiply occupies the regions occupied by each of his best candidate-constituters. Although our experience of reality shows that cats don't simultaneously occupy many quite disparate regions, that experience is silent about regions that differ as little as those occupied by Tibbles's Many.

4.3.1.2 Second objection

This second objection claims that our solution tacitly assumes that UC is false; it only follows from our proposal that Tibbles is constituted by all of the candidates because we've built it in from the start. If this is right, then our proposal is no more theoretically unified than those of Johnston and Lowe that we criticised in §4.1.

This objection fails. Although the truth of UC does block our arguments for multiple constitution, we needn't assume that it's false. We can instead remain agnostic about UC, and see whether consideration of the changes survived by cats tell for or against it. Our arguments for multiple-constitution show that these considerations tell against UC. The agnostic who seeks justification for UC in the changes that cats survive will not find one. Given our thesis that facts about constitution are explicable in terms of facts about change, it follows that UC is not true.

[11] Recall our use of 'occupy' for exact occupation: x occupies r iff x fills and fits within r.

[12] Even if several lumps can occupy a single place, the first two premisses alone conflict with our solution; for we claim that Tibbles is constituted by lumps that occupy different places at the same time.

4.3.2 Cats and maximal lumps

§1.1.4.2 considered the suggestion that Tibbles is constituted by the largest cat-like lump on his mat: the cat-like lump that includes all the cat-like lumps that include it. We rejected this because there's no reason to believe that there's a *unique* such a lump. We then considered a variant suggestion: Tibbles is the fusion of all the largest cat-like lumps on his mat. We rejected this for two reasons. Firstly, there's no reason to believe that this fusion will itself be a cat-like lump.[13] Secondly, this identifies the property of constituting a cat with the property of being a fusion of near-coincident largest cat-like lumps. But since Unger's puzzle already concerned the property of constituting a cat, this simply changes the subject.

How does this differ from our proposal? Aren't we claiming that Tibbles is constituted by the fusion of every sufficiently cat-like lump on his mat? If so, then these objections tell against our proposal too. Luckily, this isn't what we're claiming.

We claimed that Tibbles is constituted by each candidate on his mat; he is multiply constituted by all the candidates, not uniquely constituted by their fusion. Our claim is that the changes cats survive aren't fine-grained enough to distinguish between many candidates, not that they're sufficiently fine-grained as to distinguish the fusion of those candidates from everything else (other than indirectly, as the fusion of the candidates these changes don't distinguish amongst). The underlying logical point is that $R(x_1, y), \ldots, R(x_n, y)$ don't imply $R(x_1 + \ldots + x_n, y)$, where '+' denotes a fusion operation.

4.3.3 Fission and fusion

The extent of Unger's puzzle varies over time. At the level of matter, this variation is very similar to fission and fusion. But one thing we really don't want to say about fission is that the post-fission lumps both constitute one and the same object. Consider, for example, microbial reproduction, brain-transplant cases, or cutting a plant in half and re-planting the results. Two microbes, two people and two plants

[13] The following jointly imply that the fusion of Tibbles's nearby largest cat-like lumps isn't a cat-like lump: (i) they are *largest* cat-like lumps; (ii) there are many such lumps. Suppose that this fusion is cat-like. Since it includes all the other cat-like lumps, it is the unique largest cat-like lump. But this contradicts (ii).

are clearly the result, not one bi-located microbe, person or plant. Shouldn't we say the same about Unger's puzzle?

There is a significant difference of scale between the two puzzles. Fission involves much greater change than does an increase in the extent of Unger's puzzle. Although the two cases may begin in a similar manner, fission-products typically occupy non-overlapping regions and have independent futures. The changes involved in fission are too great for the object in question to survive (or so great that they only indeterminately survive or…). Variation in the extent of Unger's puzzle over time is a much smaller change than fission, one that the objects in question do seem to survive; the candidates that result don't have independent futures or occupy entirely disjoint regions of space. Despite sharing a superficially similar structure, the difference of scale between fission and Unger's puzzle justifies our treating them differently.

4.3.4 Lewis on cats and cat-constituters

Lewis rejects the distinction between cats and their constituent lumps:

> "[E]ven granted that Tibbles has many constituters, I still question whether Tibbles is the only cat present. The constituters are cat-like in size, shape, weight, inner structure, and motion. They vibrate and set the air in motion – in short, they purr (especially when you pat them). Any way a cat can be at a moment, cat-constituters also can be; anything a cat can do at a moment, cat-constituters also can do. They are all too cat-like not to be cats. Indeed, they may have unfeline pasts and futures, but that doesn't show that they are never cats; it only shows that they do not remain cats for very long." (Lewis, 1993a, p.168)

This is unpersuasive.

Firstly, possession of an unfeline past or future *does* show that something isn't a cat. Tibbles never was and will never be a scattered object, though his constituent lump probably was and will be again. So Tibbles is distinct from his constituent lump.[14]

Secondly, there are many ways a cat can be that its constituent lump(s) cannot. Cats, for example, can have the identity conditions and modal profile of a cat; but no lump can. Neither can lumps of matter purr; although matter can vibrate and set the air in motion, it's far from obvious that they, rather than the cats that they constitute, thereby come to purr. Fine (2003) provides similar examples: a statue may be Romanesque or well-made, but the bronze from which it is fashioned cannot. These categorial differences seem to be a reasonably well-entrenched feature of ordinary usage. Lewis could simply reject these differences; ordinary usage may be misleading. But this is a cost, and certainly not a theoretically neutral response to the arguments for the cat/constituter distinction.

So, there are good reasons to distinguish ordinary objects from their constituent matter. These reasons aren't unassailable, but neither are they without force. Without strong positive reason to reject these differences, which Lewis has not provided, his scepticism about the cat/cat-constituter distinction carries little weight.

4.3.5 Stating the criteria

Can we give an informative statement of either a one-level or two-level criterion of identity for cats? I certainly don't know how to. Does this tell against our proposal? This section argues that it doesn't.

4.3.5.1 Stating a two-level criterion

Applied to cats, the two-level proposal claims that a unique relation R_c satisfies:

(L2-Cat) $$\forall x \forall y (f_c(x) = f_c(y) \leftrightarrow R_c(x, y))$$

where f_c is the (partial) function from $\langle l, t \rangle$ pairs onto the cat (if any) that l constitutes at t. The question now arises: what is R_c? The only answer I can supply

[14] Lewis (1993a, pp.167–8) complains that the distinction between cats and their constituters is unparsimonious and unnecessary, given that material change can be accommodated by temporal-parts theory. But we rejected perdurance in §3.3.3.2 and the counterpart-theoretic semantics needed for stage-theory in §1.1.2.1: temporal-parts theory is an incorrect account of constitutional change.

is:[15]

> R_c is the relation that holds between $\langle l, t\rangle$ and $\langle l^*, t^*\rangle$ just in case l constitutes the same cat at t as l^* does at t^*.

There are two problems with this account of R_c. Both follow from our account of temporally relativised constitution: x constitutes y at t iff $f_c(x,t) = y$. For then our statement of (L2-Cat) amounts to:

$$\forall x \forall y (f_c(x) = f_c(y) \leftrightarrow f_c(x) = f_c(y))$$

The first problem is that this is not a two-level criterion: R_c is a relation on the $f_c(x)$s, not on the xs. The second problem is that it is obviously uninformative.

How severe are these problems? That depends on whether the two-level theorist is committed to providing an informative statement of (L2-Cat). They are not. Note first that our argument against many cats didn't require an account of R_c. Our solution to Unger's puzzle relies not on the content of R_c, but on the structural relationship between R_c, constitution and cat-identity captured by (L2-Cat).

There is no reason to expect that any single English word will be coextensive with R_c. There may (though, equally there may not) be an English disjunction, each of whose disjuncts corresponds to one (type of) instance of R_c. But the two-level view provides no reason to think that such a disjunction would be finite. Hence it provides no reason to think that we can give informative expression to R_c.

Our inability to give a non-trivial statement of (L2-Cat) reflects the relative positions of ordinary objects and matter within our cognitive architecture. We can know about and refer to matter only because we can know about and refer to the objects it constitutes: a portion of matter is accessible to us primarily as the matter of a cat, or a dog, or the top half of a trout and bottom half of a turkey, and so on. We can know that cats survive certain change in their matter only because those

[15] The difficulty in stating a two-level criterion is reinforced by the observation that the identity condition for Ks determines their existence condition also: x exists iff something is identical to x. An account in terms of the properties and relations of matter of the conditions under which a cat comes into being and all the possible changes that it can and can't survive would be a very impressive accomplishment.

changes correlate with changes in cats themselves (which we know them to survive). This doesn't refute the two-level view because (a) that view doesn't concern our cognitive, epistemic or linguistic access to cats, but the relationship between change and constitution, and (b) we can regard the trivialising account of R_c as using a relation on cats to fix the referent of an expression for a relation on matter; the semantic value of 'R_c' is fixed in just the same way as that of any other theoretical term (Lewis, 1970b).

The two-level theorist (who is not also a one-level theorist) does however regard the relative cognitive, epistemic and linguistic priorities of cats and matter as misleading. They do not reflect the underlying priority of material change over the constitution and persistence of cats. This reversal of priorities is a cost. Is there any positive argument for regarding these relative priorities as misleading, and hence also for accepting this cost? Not that I know of.

Another difficulty is that the two-level view threatens to undermine our ability to know about the persistence of cats. Our judgements about the identity and diversity of cats are based on the properties and relations of cats, not on the properties and relations of their matter. On the (pure) two-level view however, these relations bear on cat-identity only insofar as they systematically correlate with relations on the matter of their relata. Were there no such correlations, we couldn't know about the persistence of cats. A more attractive approach would more closely connect the basis on which we make judgements about cat-identity with that which determines the correctness of those judgements. This is what the one-level view offers.

We've highlighted three commitments of the two-level version of our proposal:

(i) The cognitive and linguistic priority of the persistence of cats over the properties and relations of their matter is misleading; relations on matter determine the survival of cats.

(ii) The identity condition R_c for cats is a theoretical posit, expressible only using terms for cats and constiutional vocabulary like 'the matter of...;

(iii) Our judgements about the persistence of cats are informed by a relation that bears on their persistence only indirectly, via correlations with relations on matter.

The one-level view faces none of these commitments.

4.3.5.2 Stating a one-level criterion

The one-level view fares better than the two-level view, as regards an informative account of the identity conditions for cats. Applied to cats, that proposal claims that a unique relation R_c satisfies:

(L1-Cat) $$(\forall x : Cx)(\forall y : Cy)(x = y \leftrightarrow R_c(x, y))$$

Can we express R_c in a manner that renders (L1-Cat) informative? As with the two-level criterion, I certainly don't know how to. We can however state some informative sufficient conditions for the survival of cats, thereby illuminating R_c: cats survive through the loss of their hairs, beginning and ceasing to purr, pouncing on and eating mice, and myriad other familiar kinds of change. We perceptually track cats along certain paths through space and time by tracking these kinds of change. If one of these paths connects cat x with cat y, then $x = y$. (Recall the quote from Wiggins on p.196.) Our ability to isolate these paths and to track cats along them indicates a grasp on R_c—we track cats by tracking the changes that *they* undergo, not changes in their matter or anything else—even though our language lacks the resources to express it in a manner that makes (L1-Cat) informative.[16] The idea is that we don't perceive a bare identity relation on cats, but know about identity amongst (and hence the persistence of) cats because we grasp a relation that at least approximates to R_c and which is, in typical cases, equivalent to identity amongst cats.

Given some scene-setting, we can also state some informative necessary conditions on the survival of cats, e.g.: cat x is identical to cat y only if cat x walked,

[16] Isn't a sufficient condition for the identity of cat x with cat y given in terms of paths through space and time more appropriate for a two-level criterion, than a one-level criterion? It depends on how the locations on the path are specified, and what unites those locations into a cat-survival path. The one-level theorist will conceive it as the path combining the location of cat x_1 at time t_1, with the location of cat x_2 time t_2, with.... These locations are united because the cat x_1 survived a change as a result of which it was in the location of x_2 at t_2, and x_2 survived a change as a result of which it was in the location of x_3 at t_3, and so on. The locations are thereby united by the survival of cats through change.

crawled, pounced or leaped across the room, to occupy the location that cat y now occupies. Here we need enough scene-setting to rule out human intervention, earthquakes throwing x across the room, and so on.

The one-level theorist's account of R_c is thus radically unlike the two-level theorist's. There seems little prospect for an accurate description in terms of the properties of matter of even the most simple and mundane changes that cats survive, like walking across a room. Accurate one-level descriptions however seem utterly unproblematic. The one-level theorist's difficulty with giving informative expression to R_c lies in combining these individually informative descriptions of its instances into a single statement, rather than, as with the two-level theorist, expressing them using appropriate vocabulary.

On the one-level view, we can describe, or even point to, instances of R_c. This allows us to give various informative partial accounts of that relation. We may not be able to convert this into an informative and exhaustive account of R_c, but since our account of R_c is not exhausted by its role in (L1-Cat), that principle needn't be seen as entirely trivial.

Neither the one-level nor two-level view is refuted by our inability to give an informative and exhaustive statement of the identity conditions for ordinary objects. We have however, seen that the two-level view brings costs that the one-level view does not. If we must choose between them, then, other things being equal, we should choose the one-level view. But since we needn't make a choice, the most satisfying view is probably one that combines both proposals.

4.3.6 Identity and analysis

This section considers two objections to the idea that an identity criterion is an analysis of identity. Williamson (1990, pp.144–5) presents versions of both (though he's clear that they apply only to one-level views). Hirsch (1982, ch.3 §1) presents a version of the second. We'll argue that neither objection undermines our proposal because our one-level and two-level identity criteria aren't intended as analyses of identity in other terms.

The first objection is that no analysis of identity is possible. The concept of identity is so fundamental and basic to our conceptual scheme that any attempt to analyse it in other terms is guaranteed to fail. A related complaint is that identity is utterly simple, and therefore unanalysable.

The second objection is that the notion of identity is univocal. The same identity relation holds amongst Ks as holds amongst K^*s. Different kinds don't have different identity criteria because if they did, then identity amongst Ks and amongst K^*s wouldn't be univocal.

The two-level view is immune to both complaints. The identity sign on the left of a two-level criterion is flanked by complex terms formed using functional signs and variables ranging over matter. The content of the l.h.s. therefore goes beyond bare identity. This additional complexity means that (i) the l.h.s. may be analysable, even if identity isn't, and (ii) the l.h.s. may differ in content between the criteria for Ks and K^*s, even if the notion of identity doesn't. So let's consider one-level criteria.

The first objection assumes that an identity criterion for Ks should provide an analysis of the identity relation, as it holds amongst Ks. Our one-level theorist may reject this assumption. The identity condition for Ks is supposed only to determine what changes Ks survive, not to determine what identity amongst Ks is. Furthermore, since we've granted that only a partial account of R_c may be possible, we can hardy be accused of offering an analysis of identity amongst cats. But another worry may arise: what theoretical interest do identity conditions hold, if only partial and incomplete accounts of them are possible? The answer is that they contribute to a better understanding of identity, the kind K, constitution and change by exposing some of the ways in which those concepts interact: the provision of necessary and sufficient conditions is not the only route to philosophical understanding.[17]

This also undermines the second complaint. Since our identity criteria aren't intended as analyses of identity in other terms, cross-kind variation in identity

[17] Wright (1999, §6) and Wiggins (2001, Preamble, §10) describe similar approaches to philosophical analyses of truth and identity respectively. Horsten (2010) develops a similar role for identity conditions in particular.

conditions is compatible with the univocity of identity. It follows only that different kinds of object persist through different kinds of change.

A residual worry may remain: if identity is absolutely simple, then even a partial elucidation in other terms will be impossible. The worry is baseless. One-level identity conditions aren't equivalent to identity, but to *the restriction of identity to the ordinary kind K*. Even if no elucidation of identity is possible, a partial account of the identity conditions for *K* may nonetheless contribute to our understanding of the kind *K* and its role in delimiting a domain of objects.

In short, a more holistic conception of philosophical analysis combines with the limited scope of our proposal to undermine both objections to one-level identity criteria.

4.4 Property-possession

We've argued from the priority of change over constitution to the thesis that Tibbles is constituted by each of the best candidates on his mat. Before turning to Lewis's puzzle of vague constitution, we should say a little more about this proposal. In particular, we need to address the question: what properties does Tibbles have? There is, after all, more to be said about Tibbles than what he is constituted by and when; he has a mass, colour, location and sometimes purrs.

4.4.1 Three kinds of property-inheritance

Ordinary objects inherit properties from their constituters. For example, Tibbles has a particular mass, location and colour because he's constituted by something with that mass, location and colour.

Ordinary objects don't inherit all their properties: cats purr, but lumps of matter don't; people act and think, but lumps of matter don't; statues are beautiful, Romanesque or well-made, but pieces of clay aren't. This isn't a rejection of systematic connections between these non-inherited properties of objects and those of their constitutors, just of their direct inheritance. The following discussion should be understood as restricted to inherited properties.

Were Tibbles constituted by exactly one lump, his property-inheritance would be relatively unproblematic.

Naïve Inheritance (NI) Tibbles has ϕ iff Tibbles's constituter has ϕ.

Multiple constitution makes the description 'Tibbles's constituter' improper. NI therefore fails to settle anything about Tibbles's properties. An alternative is needed. Two natural candidates are:

Universal Inheritance (UI) Tibbles has ϕ iff each of his constituters has ϕ.

Existential Inheritance (EI) Tibbles has ϕ iff at least one of his constituters has ϕ.

A third option modifies the logical form of the connection between Tibbles and his properties, by relativising it to his constituters:

Relativised Inheritance (RI) Tibbles has ϕ relative to x iff x both constitutes Tibbles and has ϕ.

We assess these in turn. UI and EI will be rejected in favour of RI. Three versions of RI will then be developed, and two defended.

4.4.2 Against Universal Inheritance

Suppose that Tibbles's constituters don't all have exactly the same mass. UI implies that Tibbles doesn't have any particular mass, despite being constituted by lumps that do. Likewise for spatial location. But then in what sense is he a material object?

One might respond that although Tibbles doesn't have any particular mass or location, he is massive and he is located because each of his constituters is massive and located. But surely it's analytic that something is massive or located only if it has some particular mass or particular location. It is obscure what being massive or located might amount to, if not the possession of some particular mass or location. So we should reject this response.

As well as undermining Tibbles's status as a material object, UI undermines our ability to know about him. How can we causally interact with something that doesn't have a spatial location or mass? And if we can't causally interact with Tibbles, then it's unclear how we can know about him. Although we can causally interact with his cat-like constituters, none of those constituters is Tibbles. So causal interaction with them doesn't alleviate the problem of how we can know about Tibbles.

In light of these two problems, we should reject UI.

4.4.3 Against Existential Inheritance

Suppose that Tibbles's constituters don't all have exactly the same mass. EI implies that Tibbles has many masses. But since distinct masses are incompatible determinates of the same determinable, this is impossible. At best, EI requires substantial modifications to our ordinary conception of property-incompatibility and the determinate/determinable contrast. We therefore reject it.

4.4.4 In defence of Relativised Inheritance

Only RI remains. On this view, Tibbles has a range of particular masses and locations, each relative to one his constituters. So unlike UI, RI doesn't deprive Tibbles of having a particular mass and colour. And since Tibbles doesn't have incompatible masses *simpliciter*, but only relative to different constituters, RI, unlike EI, doesn't conflict with our ordinary conception of property-incompatibility.

An argument of sorts from our core thesis to RI is possible. According to our proposal, Unger's puzzle shows that, when examined closely enough, the changes Tibbles survives supply him with a branching path through space and time. No single branch contains the whole of his history, though each is one of his histories. So when Tibbles's inherited properties differ across branches, no single assignment of those properties to him can tell the whole story. In order to say everything there is to say about Tibbles's properties, we have to say what branch they are found on. RI implements this idea.

This section elaborates RI by defending it against an objection. Two defensible forms of the view will be found, though one is preferable to the other.

4.4.4.1 A problem for RI

§§4.4.4.2–4.4.4.4 examine three responses to the following objection to RI: according to RI, Tibbles has few, if any, intrinsic properties; in particular, he has no mass, location, shape, or any other intrinsic property that he should inherit from his constituters.

Without supporting argument, we should be unmoved by this objection. Following is one such argument.

According to RI, the sense in which Tibbles has a mass differs from the sense in which a lump has a mass; for Tibbles has a mass relative to a lump, while lumps simply have mass. Tibbles's having a mass therefore involves his bearing a relation to things with that mass, while a lump's having a mass does not. We can now adapt an argument of Lewis's:

> "I protest that there is... nothing in the picture that has [*5kg*] *simpliciter*. ... Instead of having [*5kg*] *simpliciter*, [Tibbles] bears the [*having-relative-to*] relation to it and [a lump l]. But it is one thing to have a property, it is something else to bear some relation to it. If a relation stands between you and your properties, you are alienated from them." (Lewis, 2002, p.5. Lewis is objecting to theories that explain intrinsic change by treating instantiation as a relation between objects, properties and times. We've modified his example to fit the present case.)

There are two claims here:

(6) *5kg* is an intrinsic property of Tibbles.

(7) Having an intrinsic property is not a matter of bearing a relation to it.

We've already got:

(8) If Tibbles has ϕ only relative to a constituter, then his having ϕ is a matter of his bearing a relation to ϕ (and that constituter).

Together these imply that Tibbles doesn't have *5kg* only relative a constituter. Generalising: ordinary objects don't possess their intrinsic properties only relative to their constituters. So RI is false.

How might we respond? Premisses like (6) concern the paradigms that fix the content of our notion of intrinsicality. The goal is a theory of properties and intrinsicality that is compatible with such claims. So the defender of RI must reject (7) or (8). Two kinds of resistance to (7) are possible:

> Having any property is a matter of bearing a relation to it.
>
> Although having some properties is not a matter of bearing a relation to them, the having of inherited properties by ordinary objects is.

These are examined in §4.4.4.2 and §4.4.4.3 respectively. Only the latter is defensible. §4.4.4.4 presents an attack on premiss (8):

> Having a property relative to a constituter need not be a matter of bearing a relation to that property (and constituter).

Although the second attack on (7) is defensible, we'll see that this third option is preferable.

4.4.4.2 An instantiation relation?

This section considers rejecting (7) on the grounds that all property-possession is analysable into the bearing of a relation between object and property. Bradley's Regress shows that this view is not tenable: some structures cannot be analysed into the bearing of a relation amongst their constituents.[18]

To see this, consider the thesis:

Relational Analysis of Instantiation (RAI) Whenever an object x_1 bears a relation R to objects $x_2, \ldots, x_n$, this is analysable into the bearing of an instantiation relation I amongst $R, x_1, \ldots, x_n$.[19]

Suppose that a_1 bears R to $a_2, \ldots, a_n$. By RAI: this is analysable into the bearing of an instantiation relation I amongst $R, a_1, \ldots, a_n$. By RAI: this last fact is analysable into the bearing of an instantiation relation I' amongst $I, R, a_1, \ldots, a_n$. By RAI: this is itself analysable into the bearing of an instantiation relation I'' amongst $I'I, R, a_1, \ldots, a_n$. And so on *ad infinitum*. The output of each analysis by RAI is a suitable input for analysis via RAI: each relational fact is the first element of an

[18] Our mention of structures doesn't bring commitment to entities that are structures. There may be no structures, but only entities that are (collectively) structured. Talk about structures is for convenience only.

[19] Treat monadic properties as the limiting case of relations.

infinite series of further relational facts; the obtaining of any fact in this series is analysed into the obtaining of its successor.

Two problems arise, one metaphysical and one epistemological. The metaphysical problem assumes that, in a good analysis, the analysans is more metaphysically fundamental than the analysandum. RAI then implies that (a) every relational fact decomposes into an infinite series of ever more fundamental relational facts. If the defender of RAI finds series' of this kind problematic, then they might respond by excluding some elements of the series from the domain of RAI. This implies that (b) there is a seemingly arbitrary point in the series at which the analysis terminates. Let us grant for the sake of argument that the notion of x being more metaphysically fundamental than y—conversely: y ontologically depending on x—is contentful. Even given this dubious assumption, neither (a) nor (b) is obviously problematic. Hence Bradley's Regress is not obviously a metaphysical problem.

Consider (a). Although it *sounds* problematic, it is unclear what *argument* might be brought against infinite sequences of ever more fundamental facts. This isn't to say that such sequences aren't problematic, only that it's unclear what could justify a view either way.

Consider (b). What is the relevant notion of arbitrariness? Well, the terminus of our series of ever more fundamental relational facts would presumably be utterly fundamental: its obtaining cannot be analysed into the obtaining of any further relational fact. So there is no account in other terms of why such a series terminates where it does. So the existence of a terminus in our series of relational facts is not itself objectionable. And neither is the belief that there is such a terminus, given an argument against infinite chains of ontological dependence. It would be arbitrary to believe, of any element in the series, that it is the terminus. But since that isn't what (b) requires, it's no reason to find (b) objectionable. It's unclear what further argument for a problematic form of arbitrariness might be appealed to.

Since neither (a) nor (b) is obviously problematic, RAI's relational analysis of instantiation is not obviously a metaphysical problem. It is better seen as an epistemological problem. The following draws on Fraser MacBride's (2005a) discussion, though it's unclear whether he would agree with our conclusion.

The (a?) point of analysis is explanation. One ought to endorse an analysis therefore only if one has good reason to believe that the analysans can explain the analysandum. One ought to endorse RAI therefore, only if one has good reason to believe that the bearing of a relation I amongst $R, x_1, \ldots, x_n$ can explain R's holding amongst $x_1, \ldots, x_n$. But since RAI applies also to I's holding amongst $R, x_1, \ldots, x_n$, one must also have good reason to believe that the holding of a relation I' amongst $I, R, x_1, \ldots, x_n$ can explain this fact. And since RAI applies to I's holding amongst $I, R, x_1, \ldots, x_n$, one must also have good reason to believe that.... At each stage, the phenomenon being explained is the very phenomenon used in the explanation: the bearing of a relation. One therefore ought not to endorse RAI unless one already had good reason to believe RAI successful; RAI should not come to be accepted by those who don't already believe it. In the absence of such prior belief, we should reject RAI. Bradley's Regress shows that one can never acquire reason to believe RAI unless one already has reason to believe it.

Since we shouldn't believe RAI, we shouldn't believe that entering into the relational structure $R(a, b)$ by its constituents R, a, and b is analysable into the bearing of a relation amongst those constituents. Some structure is not analysable into the bearing of relations.

The situation resembles an argument with the global sceptic. The sceptic's position is consistent and very difficult, if not impossible, to refute. The question is whether we should join them in that position. Unless we are already sceptics, there seems no reason to do so. And unless we are already committed to RAI, there seems no reason to believe it, or even to believe that the kind of explanation it posits could be successful.

If this is correct, then there is a difference between possessing a property and bearing a relation to it. If Tibbles "has" *5kg* either by being related to it, or by being related to something else that has it, then he does not really have *5kg* (or, at least, does so only in an extended sense defined in terms of the primary sense). The next section considers a different objection to (7) that grants this conclusion, but denies that this difference is problematic.

4.4.4.3 A relational account of inheritance

Let us grant the conclusion of the previous section: if Tibbles has a mass-property only relative to a constituter, then he has that property only in an extended sense, not the primary sense in which his constituter has it. Is this an objectionable difference in sense? If not, then we may reject thesis (7) from p.226, and the argument against Relativised Inheritance along with it.

Here is one flawed reason to think that this difference is not objectionable. When Tibbles has *5kg* in the extended sense of being related to something that has *5kg* in the primary sense, the relation in question is *constitution*: Tibbles is constituted, or made out of, something that really does have *5kg*. The flaw is that it's unclear why this should help.

The concern about RI was that if it is true, then Tibbles does not really have any mass-property. Since Tibbles is distinct from each of his constituters, what mass-properties they have is irrelevant to this concern. This is not alleviated by calling the relation between Tibbles and a lump 'constitution'. One reason is that we are using 'constitution' only as a place-holder for whichever relation occupies the pre-theoretic *making-up* role (§4.2.6). There is a more powerful reason also. If constitution brought elimination or reduction, then this would alleviate the problem. But it does not: cats are neither reducible to, nor eliminable in favour of, their matter. What's needed is some kind of metaphysically substantial notion of constitution, on which constituted and constituter, although distinct, are less than completely distinct; a halfway house between elimination and non-elimination.[20] It is far from clear that there is any such notion. And even if there is, the position that results is in tension with our claim that change is prior to constitution: an object's underlying nature is not given by listing its microscopic constituents and the way in which they are put together, but by specification of the changes it survives. So let us set this view aside.

A more promising strategy asks why we should deny that Tibbles's having *5kg* is a matter of his being related to a constituter that has *5kg* in the primary sense. Why

[20] Claims like the following only add to the obscurity: constituted entities are no increase in being beyond, or an ontological free lunch given, the existence of their constituters.

should that be objectionable? Lewis's work on change provides the only answer that I know of.

How is intrinsic change possible? Nothing can be both bent and straight; yet many things are bent at one time and straight at another. What role do the times play here? One answer is that intrinsic properties like *bent* and *straight* are really relations to times; they are really relations, *bent-at-* and *straight-at-*, between objects and times. Against this view, Lewis claims:

> "As we persist, we change. And not just in extrinsic ways, as when a child was born elsewhere and I became an uncle. We also change in our own intrinsic character, in the way we ourselves are, apart from our relationship to anything else.... When I change my shape, that isn't a matter of my changing relationship to other things, or my relationship to other changing things. *I* do the changing, all by myself." (Lewis, 1988a, p.187)

But if Lewis's having a shape is a matter of his being constituted by something with that shape, then his changing shape is a matter of his changing relations to other changing things. So, Lewis will conclude, his having an intrinsic property like a shape or a mass is not a matter of his being constituted by something with that shape or mass.

MacBride (2001, §2) notes that Lewis's argument fails. A change in the shape of a temporally located object is not a matter of its changing relations to other changing temporally located things. But it doesn't follow that this change isn't a matter of the object's changing relations to times, to temporal locations themselves. This is simply under-determined by our experience of change.

Likewise, a change in the shape or mass of a temporally located object is not a matter of its changing relations to other temporally located things *that don't constitute it*. But it doesn't follow that this change isn't a matter of the object's changing relations to other temporally located things that do constitute it. This too, is under-determined by our experience of change. In fact, it seems reasonable to think that a change in an object's mass is a certain kind of change in the matter that constitutes it.

I know of no other reason to deny that Tibbles's having a mass is a matter of his being constituted by something with that mass. We are therefore free to maintain that it is, and hence to reject premiss (7) of the argument against Relativised Inheritance. But a sense of unease remains; for there is an important sense in which Tibbles does not possess any of the properties inherited from his constitutors, viz. the primary sense in which they possess those properties. Although we've found no reason to reject this view, it doesn't follow that we should accept it. We may still hope for better. That is what the next section seeks to provide.

4.4.4.4 Non-relational Relativised Inheritance

This section examines a rejection of premiss (8) of the argument against Relativised Inheritance. According to such views, this:

Tibbles has ϕ relative to l.

doesn't imply this:

Tibbles has ϕ by bearing a relation to ϕ and l.

We've seen that an object's having a property (or standing in a relation) needn't be analysable into the bearing of a relation between object and property (or between relata and relation). Let S be a structure with constituents $s_1, \ldots, s_n$. Then we've seen that S needn't be analysable into the bearing of a relation amongst $s_1, \ldots, s_n$. Not all structure is relational structure. This section argues that the structures posited by Relativised Inheritance needn't be relational structures.

Consider the structure that Tibbles, *5kg* and l enter into when Tibbles has *5kg* relative to l. We'll call this structure S. We want to know: is S analysable into the bearing of a relation between Tibbles, *5kg* and l? I can see three unpersuasive reasons to think that it is.

The first reason to endorse the relational analysis of S begins with the following necessary equivalence:

Tibbles has *5kg* relative to l iff: l has *5kg* and l constitutes Tibbles.

Necessary equivalence makes it formally permissible to posit an analysis. But it doesn't mandate it. Consider the following necessary equivalences:

Socrates exists iff {Socrates} exists.

Grass is green or grass is not green iff $2 + 2 = 4$.

The r.h.s. of the former certainly doesn't analyse the left; and the left is an, at best, controversial analysis of the right. Neither side of the the second is plausibly any kind of analysis of the other. Necessary equivalence therefore doesn't imply analysis.

The second reason appeals to the principle:

> Every structure with more than two constituents is analysable into the bearing of a relation amongst those constituents.

Since S has three constituents, this implies that S is analysable into the bearing of a relation amongst them. Unfortunately for the defender of this second argument, this principle is false. When a relation holds between two relata, the resulting structure has three constituents. But we've already seen that not all such structures are analysable into the bearing of further relations amongst their constituents.

The third and final reason appeals to two theses:

(i) S has three constituents.

(ii) An instantiation of a monadic property ϕ is a structure with (at most) two constituents: ϕ itself and the object that has ϕ.

It follows that S is not an instantiation of a monadic property: either *5kg* is really a relation, or S is analysable into the bearing of a relation amongst its constituents.

This argument fails because (ii) is false: the number of entities involved in an instantiation is no guide to the degree of the property instantiated. (for discussion see MacBride, 2005b, §2.) Two examples to make the general point: *identity* is a dyadic relation whose instantiations can only involve one other entity; *at least as large as* is a dyadic relation some of whose instantiations involve two other entities, and others of which involve only one. Three examples to make the point for the special case of monadic properties: *arranged in a circle* is a monadic (plural collective) property whose instantiations involve different numbers of entities on different occasions (furthermore, those instantiations never involve only one other entity); *being a property* and *being self-instantiating* are properties some of who's instantiation feature only one constituent, namely that property itself. Since the number of objects involved in an instantiation alongside the property or relation instantiated is no guide to adicity, thesis (ii) above is false.

Adicity is not a feature of how many objects are involved in instantiations, but of the kinds of resemblance and difference marked by a property or relation. Our thesis of Relativised Inheritance concerns the number of entities involved in instantiations of inherited properties, not varieties of resemblance and difference. That thesis is therefore compatible with the monadicity of *5kg* and the claim that S isn't analysable into the bearing of a relation between Tibbles, *5kg* and l.

I can find no other argument for premiss (8) from p.226. We are therefore free to reject it, and the argument against Relativised Inheritance along with it.

4.4.5 Property-possession: concluding remarks

This section began with three accounts of property-inheritance. We settled on the relativisation of inherited properties to constituters and saw two defensible versions of this view. One analyses x's possession of ϕ by l into: l possesses ϕ and constitutes x. The other rejects that analysis: relativised instantiational structure is a *sui generis* variety of structure, not analysable in other terms. Although both views are defensible, the second carries an advantage over the first: the connection between Tibbles and his inherited properties is not mediated by any relation. We turn now to Lewis's puzzle of constitutional vagueness. One way of developing our proposal (discussed in §4.5.5) in order to accommodate this puzzle will favour the first form of Relativised Inheritance over the second.

4.5 Vagueness

Up to now, we've focused on Unger's puzzle of too many candidates and assumed that it's clear which the best candidates are. This section relaxes this assumption and extends our proposal to Lewis's puzzle of vague constitution and borderline candidates.

§4.5.1 begins by arguing from an unmodified version of our proposal to unclarity in constitution, mereology and inherited properties. §4.5.2 then argues that despite this unclarity, our proposal is incompatible with the Sharpening View of constitutional vagueness, unless it is modified in some way. We can draw two conclusions from this. Firstly, Unger's puzzle is a source of unclarity in constitution, but that unclarity is not a form of vagueness: Unger's puzzle of too many best candidates is not a puzzle of full-blown vagueness, though it does give rise to a more limited form of unclarity. Secondly, our proposal must be modified in order to accommodate constitutional vagueness. Different kinds of response to the initial argument for incompatibility provide different strategies for extending our proposal to constitutional vagueness. To this end, we close §4.5.2 by uncovering a hidden assumption within that argument and three potential responses. These responses are developed in §§4.5.3–4.5.5.

4.5.1 Unclarity in constitution, inheritance and parthood

We've proposed that Tibbles is constituted by each of his Many, and that he has his inherited properties only relative to these constituters. This section argues from both theses to unclarity in constitution, inherited properties and parthood. The next section argues that this is not a form of vagueness, but a more limited variety of unclarity. There is work to be done before we can accommodate genuine vagueness.

4.5.1.1 Unclarity in constitution

'Constitution' was introduced as a name for whichever relation occupies the pretheoretic "making up" role. Although this role is implicitly defined by our use of language as a whole, there is some leeway about just how closely the constitution-relation fits this role: a reasonably good fit may be good enough. Our thesis of multiple-constitution brings deviation from this role. Two examples: we may ordinarily talk about *the* wool from which a jumper is made; or we might say, in an only slightly more theoretical vein, that a statue is in the same place as *the* particles that make it up. These definite descriptions manifest the assumption that constitution is unique, thereby incorporating that assumption into the implicit definition of the constitution-role.

A question now arises: how does the extension of ordinary constitutional vocabulary relate to our postulated relation of multiple-constitution? To avoid confusion, we'll call this relation CON and write as if 'constitutes' were part of ordinary English. We'll also restrict these notions to a single kind of ordinary object, specification of which we'll tend to leave tacit. Then our question is: what is the relation between CON and the extension of 'constitutes'? The puzzle is that we use 'constitutes' as if it were one-one, while our proposal is that CON is many-one.

Let us treat relations as sets of ordered pairs, and ignore non-constitutional vocabulary for simplicity. Then we can frame two hypotheses about the relationship between 'constitutes' and CON:

Identity An interpretation s is intended iff $[\![\text{'constitutes'}]\!]_s = \textsc{con}$.

Inclusion An interpretation s is intended iff:

(i) $[\![\text{'constitutes'}]\!]_s \subseteq \textsc{con}$; and

(ii) If $\langle x, z\rangle \in [\![\text{'constitutes'}]\!]_s$ and $\langle y, z\rangle \in [\![\text{'constitutes'}]\!]_s$, then $x = y$; and

(iii) If $\langle x, z\rangle \in \textsc{con}$, then, for some $y : \langle y, z\rangle \in [\![\text{'constitutes'}]\!]_s$.

According to Identity, there is exactly one intended interpretation of 'constitutes': the relation CON. Since CON is many-one, descriptions like 'the clay that makes up the statue' are improper. Sentences featuring them are therefore be untrue.

According to Inclusion, there are many intended interpretations of 'constitutes'. Condition (i) ensures that constitutional vocabulary aims at describing the CON facts: if x constitutes y, then $\textsc{con}(x, y)$. Condition (ii) ensures that constitution is one-one. Condition (iii) ensures that everything that should have a constituter—everything to which something bears CON—does have a constituter. (i)–(iii) together ensure that the intended interpretations of 'constitutes' are the minimal deviations from CON to give a relation that fits the constitution-role. 'Constitutes' has many intended interpretations because there are many such minimal deviations. This brings unclarity in constitution without making 'the matter of Tibbles' improper.

Two semantic pressures must be reconciled. The first is our use of constitutional vocabulary to describe the object-matter relation CON. The second is our use of constitutional vocabulary as if constitution were unique. Our proposal brings these pressures into conflict by making CON many-one. The Identity hypothesis resolves in favour of the first pressure. This results in improper descriptions and untruth. The Inclusion hypothesis resolves in favour of the second pressure. This results in many intended interpretations of 'constitutes', and hence unclarity in constitution.

Two arguments suggest that the second kind of resolution wins. The first is a methodological argument: it makes true a greater proportion of ordinary talk, and theories that do so are *ceteris paribus* preferable to theories that don't. The second is a metaphysical argument: semantic values are (to a significant extent) determined by which sentences ordinary speakers hold true. Our solution to Unger's puzzle therefore entails unclarity in ordinary constitutional vocabulary.

4.5.1.2 Unclarity in inheritance

This section argues from our proposal to unclarity in ascriptions of inherited properties. This unclarity will infect mass, location, shape, and any other property that can vary across Tibbles's constituters (provided it is inherited by cats). Given the following connection between property-ascription and predication, this unclarity will extend beyond explicit property-ascriptions:

> x has the property of being F iff x is F.

Ordinary property-ascription isn't relativised to a constituter. Since our proposal relativises inherited properties to constituters, the following challenge arises: to convert the relativised ascriptions into truth-conditions for un-relativised ascriptions.

An un-relativised notion $R(x)$ is most naturally obtained from a relativised one $R^*(x,y)$ by closing the y-position in R^*. We might use a quantifier or other variable-binding operator; or we might substitute y for a singular term. We'll opt for this second alternative. Each intended interpretation s selects a lump l_s from amongst Tibbles's constituters and supplies the following truth-condition to property-ascriptions:

'Tibbles has ϕ' is s-true iff Tibbles has ϕ relative to l_s.

Since Tibbles has many constituters, many intended interpretations of property-ascription result. Since Tibbles's constituters have different masses, the truth-values of ascriptions of inherited properties to Tibbles vary across intended interpretations. The result is unclarity about Tibbles's mass. Likewise for other inherited properties.

Although other accounts of property-ascription are possible, this one is the best; for only it can respect the truth of:

Tibbles has the same location and mass as his constituter.

This is secured by the following penumbral connection:

x s-satisfies 'constitutes Tibbles' iff: Tibbles has ϕ is s-true iff Tibbles has ϕ relative to x.

This connection ensures that the lump that features in the s-truth-conditions for property-ascriptions is the lump that s counts as Tibbles's constituter. Unclarity in constitution induces unclarity in inherited property-ascription via analytic connections between constitution and inheritance.

4.5.1.3 Unclarity in Parthood

Finally, we can argue from unclarity in constitution to unclarity in the mereology of ordinary objects.

On one kind of view, object-mereology is definable via constitution and a parthood relation on matter:

x is part of an ordinary object o iff x is part of the lump that constitutes o.

Given this, 'is part of' has many intended interpretations if 'constitutes' does: unclarity in constitution entails unclarity in object-mereology.

This kind of view is controversial. Although my heart is part of me, it doesn't seem to be part of any lump of matter; my heart's matter is part of my matter, but my heart itself isn't. Me and my heart are thus a counterexample to the equivalence above. Approaches to object-mereology based around that equivalence are incompatible with hierarchical conceptions on which my heart isn't part of any lump of matter.[21] Still, parthood and constitution are connected:

> If x is part of o, then x's constituter is part of o's constituter.

An object's parts must be constituted by parts of its matter. Since different sharpenings count different candidates as Tibbles's constituter the result is unclarity about Tibbles's microscopic parts. Unclarity in constitution entails unclarity in object-mereology.

4.5.2 Constitutional vagueness?

This section argues that despite the argument from our proposal to unclarity in constitutional vocabulary, there is work to be done before we can accommodate constitutional vagueness.

§4.5.2.1 begins by arguing that our proposal is, as it stands, incompatible with a Sharpening-theoretic account of constitutional vagueness. This shows that the unclarity argued for in the previous section is a more limited phenomenon than genuine vagueness. Our proposal must be extended somehow in order to accommodate constitutional vagueness. We'll organise our discussion of such extensions around different kinds of response to the initial argument for incompatibility. To this end §4.5.2.2 identifies a hidden assumption behind that argument, and §4.5.2.3 sketches three kinds of response. These are investigated in §§4.5.3–4.5.5. The first response modifies neither our proposal nor the Sharpening View, but rejects the hidden assumption. The second retains the assumption and the Sharpening View without modifying our proposal, but explains away the appearance of constitutional vagueness. (A notational variant on this second response endorses an account of constitutional vagueness other than the Sharpening View.) The third

[21] We might respond with an alternative account of object-mereology: x is part of o iff the matter of x is part of the matter of o. If this is acceptable to the defender of the hierarchical view, the argument from unclarity in constitution to unclarity in object-mereology goes through as before. Difficulties remain however: the sleeve of a woolen jumper is not part of the thread from which the jumper is made, though the matter of the sleeve is part of the matter of the thread. The following argument in the text shows that we can sidestep these issues.

retains the assumption and the Sharpening View, but modifies our proposal by allowing gradual constitution and hence (what the Sharpening theorist regards as genuine) constitutional vagueness.

4.5.2.1 Sorites-susceptibility and higher-order borderline cases

This section argues that our proposal requires modification in order to accommodate the Sharpening View's conception of constitutional vagueness. We proceed by arguing that, as it currently stands, our proposal is incompatible with that conception. Subsequent sections examine potential modifications of our proposal to avoid this argument, and hence to accommodate constitutional vagueness.

On our proposal, many lumps bear the relation CON to Tibbles. §4.5.1 argued that each of these lumps a borderline case of a cat-constituter. Consider a Sorites series S such that (i) S begins with lumps that clearly don't constitute Tibbles, (ii) S terminates with lumps that bear CON to Tibbles, and (iii) for each element x of the series, its successor x' differs only very marginally from x in respects relevant to bearing CON to Tibbles. The following Sorites principle is intuitively plausible:

(9) $$\forall x(\neg x \text{ constitutes a cat } \rightarrow \neg x' \text{ constitutes a cat})$$

The Sharpening theorist offers two explanations for why these Sorites principles are attractive, despite their being provably false (§2.5.2). We now argue that our proposal is (as it currently stands) incompatible with both these explanations, and hence incompatible with a Sharpening-theoretic account of the Sorites-susceptibility of 'constitutes'.

The Sharpening theorist's first explanation for the attraction of (9) appeals to their conception of vagueness as the result of imposing a non-gradual classification onto a gradual transition. When we do so, nearby points in the transition will differ very little both in those respects involved in the transition, and in their relations to our linguistic behaviour. A Sorites principle is a natural, though incorrect, way of articulating this.

Our proposal can't accommodate this first explanation because CON is non-gradual. Adjacent candidates in our Sorites series S can (and will) differ significantly w.r.t. bearing CON to Tibbles. (§4.5.5 modifies our proposal to allow that CON is gradual.)

The Sharpening theorist's second explanation for the attraction of (9) begins by observing that no instantiation of its negation is assertable because each is borderline. The Sharpening theorist then attributes to typical speakers a mistaken slip from the unassertability of these instantiations to their falsity, and thereby to the truth of (9).

On the face of it, this first explanation should be applicable to (9). Instantiations of its negation are of the form:

$$\neg a \text{ constitutes a cat} \wedge a' \text{ constitutes a cat}$$

§4.5.1 argued that the best candidates (the ones that bear CON to Tibbles) are all borderline cat-constituters. So each sentence of the above form is either clearly false or borderline, and hence unassertable. So the Sharpening theorist's attribution of a mistaken slip from unassertability to falsity should explain the attraction of (9).

Matters are not quite so clear-cut. The present case is relevantly unlike a typical Sorites. Although the series *S* terminates with the best candidates to be cat-constituters, none of them is a clear cat-constituter (despite it being clear that one of them is a cat-constituter). Since these are the best cases and each has everything that could be desired in order to be a case—each bears CON to Tibbles—they might quite easily be mistaken for clear cases (as the reasoning behind Unger's puzzle suggests that they are). Indeed, they only fail to be clear cases because of the conflicting semantic pressures governing constitutional vocabulary that we described in §4.5.1. Unlike typical cases of vagueness, they aren't borderline cases because our use of language privileges no one classificatory boundary, but because two features of that use conflict. If this is right, then when a is the last non-candidate, there's a sense in which 'a' constitutes a cat' is clearly true; the sense in which our use of 'constitutes' aims at CON. Since '$\neg a$ constitutes a cat' is then also clearly true, the result is a clearly true instance of the form displayed above. The Sharpening View's explanation for the attraction of Sorites principles therefore doesn't seem to extend to principle (9).

These arguments show that there's work to be done before our proposal can accommodate a Sharpening-theoretic account of the Sorites-susceptibility of 'constitutes'. There's also work to be done before we can accommodate a Sharpening-theoretic account of higher-order vagueness in 'constitutes'.

§2.9.9 developed an account of higher-order vagueness that appeals to metasemantic gradualness: a series of interpretations, each of which fits the meaning-determining facts only slightly less well than its predecessor, gives rise to many intended interpretations of 'intended interpretation', and many intended interpretations of 'intended interpretation of 'intended interpretation' ', and so on. But consider a sharpening *s* on which the last non-candidate in the series satisfies 'constitutes Tibbles': the *s*-extension of 'constitutes' isn't a sub-relation of CON, in violation of our first condition on intended interpretations of 'constitutes' (p.236). Other things being equal, *s* therefore fits the meaning-determining facts significantly less well than interpretations on which the extension of 'constitutes' is a sub-relation of CON. This limit on metasemantic gradualness prevents any second-order borderline cases from separating the clear non-cases from the first-order borderline cases. Our proposal therefore needs modifying before it can accommodate the Sharpening theorist's account of higher-order constitutional vagueness.

§§4.5.3–4.5.5 generalise our proposal to accommodate constitutional vagueness. They do so by examining three kinds of response to these arguments. The next section begins by identifying a hidden assumption on which these arguments rely.

4.5.2.2 A hidden assumption

This section identifies a hidden assumption in the previous section's argument for the incompatibility of our proposal with a Sharpening-theoretic account of constitutional vagueness. The assumption is a conception of content-determination akin to that of Lewis (1983a, 1984).

According to Lewis, content is determined by (at least) two features of interpretations: how well they fit our linguistic behaviour, and the Eligibility of the semantic values they assign to our vocabulary. Eligibility is a measure of intrinsic suitability to be meant. Typically, and certainly in Lewis's view, the Eligibility-ordering is identified with the naturalness-ordering.[22]

[22] To accommodate singular terms, the naturalness-ordering on properties needs extending to an

Consider the problem facing the Sharpening theorist's first account of the Sorites principle (9). The problem was that since CON is non-gradual, some adjacent candidates in the Sorites series S will differ significantly in respects relevant to whether they satisfy 'constitutes Tibbles': one but not the other will bear CON to Tibbles. Why should this matter to how well suited those candidates are to satisfy 'constitutes Tibbles'? Those candidates won't differ significantly in their relations to our use of constitutional vocabulary. The only alternative answer seems to be that interpretations that make the extension of 'constitutes' a sub-relation of CON are *ceteris paribus* more Eligible than those that don't.

Consider the problem facing the Sharpening theorist's second account of the Sorites. The problem was that since the distinction between those candidates that bear CON to Tibbles and those that don't isn't vague, and hence a significant sense in which there are no borderline candidates. This provides a significant sense in which some sentence of the following form is clearly true:

$$\neg a \text{ constitutes Tibbles} \wedge a' \text{ constitutes Tibbles.}$$

Unless interpretations on which the s-satisfier of 'constitutes Tibbles' bears CON to Tibbles fit the meaning-determining facts significantly better, *ceteris paribus*, than those on which it doesn't, this argument fails. It's unclear what could justify that, other than appeal to the Eligibility of CON.

Finally, consider the problem for the Sharpening theorist's account of higher-order vagueness. This assumed that the following suffices, *ceteris paribus*, for a significant difference w.r.t. how well interpretations s, t fit the meaning-determining facts: the s-extension of 'constitutes' is a sub-relation of CON, though the t-extension of 'constitutes' isn't. What justifies this, if not the Lewisian conception of content-determination? Such differences seem insignificant w.r.t. fit with use. Appeal to a significant difference w.r.t. Eligibility seems to be the only alternative.

ordering on objects. Lewis (1983a, p.49) suggests that we do so by appeal to how well their boundaries are demarcated by natural properties.

4.5.2.3 Three kinds of response

We've seen that the Lewisian conception of content-determination is assumed by the argument in §4.5.2.1 for the incompatibility of our proposal with the Sharpening View's account of constitutional vagueness. This section distinguishes three kinds of resistance to that argument. Each provides one way of extending our proposal to accommodate constitutional vagueness.[23]

(i) Reject the Lewisian conception of content-determination.

(ii) Deny that constitution is vague, and explain away the appearance that it is.

(iii) Develop an account of gradual constitution.

Response (i) retains our proposal and the Sharpening View without modification; the arguments to show that some modification is needed are rejected instead. Response (ii) accepts the arguments for the incompatibility of our proposal with constitutional vagueness, taking them to show that constitutional vagueness is impossible. Advocates of this response must explain away the appearance of constitutional vagueness in a way that doesn't extend to all other cases of (apparent) vagueness, and thereby undermine the Sharpening View. Response (iii) also accepts that the problems are genuine, but takes them to show instead that our solution to Unger's puzzle will not do as it stands; the goal is to make CON more like the properties and relations relevant to typical cases of vagueness. The following sections consider these in turn. We won't come to a settled view about which is preferable; each is defensible, though each has its costs.

4.5.3 Content-determination without Eligibility

This section presents an account of constitutional vagueness that rejects Lewis's Eligibility-based conception of content-determination, leaving the Sharpening View and our response to Lewis's puzzle unmodified. We begin with some concerns about the Lewisian account of content-determination.

[23] A fourth option isn't considered here: reject the Sharpening View of vagueness wholesale in favour of an alternative.

4.5.3.1 Against Eligibility

Although popular, the Lewisian conception of content-determination is somewhat mysterious. Its best, and probably only, motivation is to respond to Kripke and Putnam's arguments for scepticism about meaning (Kripke, 1982; Putnam, 1980). If there are other, better solutions, then the view is unmotivated. We can't investigate the alternatives here, but it's worth noting this way in which Lewis's view is a hostage to theoretical fortune.

The view comes in two varieties, depending on whether Eligibility is identified with naturalness or not. We'll raise some worries about both varieties.

Consider the view that identifies Eligibility with naturalness. We should ask: why does this connection hold? Why is a more natural property a better candidate semantic value than a less natural one, other things being equal? As Lewis (1983a, pp.54–5) makes clear, the answer isn't that we intend to use language to mark reasonably natural distinctions. Not only is it highly dubious that we have such intentions, but that answer presupposes an account of content-determination for intentions; yet the arguments for meaning-scepticism apply to thought and language both. No alternative account of the Eligibility-naturalness connection is forthcoming. The result is a surprising and unexplained connection between a property's naturalness and its suitability to be expressed by a predicate. This connection comes not from an investigation into the nature of meaning, but a desire to block a problematic argument. Maybe we should hope for no more than this, but it is hard to see the result as a unified theoretical package.

Consider now the view that distinguishes Eligibility from naturalness. So what is Eligibility? There seems no alternative independently motivated ordering on candidate semantic values whose identification with Eligibility would be any less mysterious than that of naturalness. So the Eligibility-ordering must be taken as a *sui generis* kind of semantic fact: some potential meanings are just better meanings than others. On the one hand, this doesn't address Kripke and Putnam's sceptical challenges, so much as simply insist that there is a response to them. On the other hand, it blocks an account of semantic facts in broadly naturalistic or physicalistic terms.

We haven't shown that Lewis's Eligibility-based account of content-determination is false. We have shown however, that its motivation is tenuous and it either (i) brings mysterious connections between seemingly disparate kinds of fact, or (ii) blocks a naturalistic account of semantics. These are good reasons to be sceptical about it. And once that scepticism is in place, we should also be sceptical of the arguments purporting to show that our proposal needs modifying in order to accommodate a Sharpening-theoretic account of constitutional vagueness.

4.5.3.2 Constitutional vagueness without Eligibility

Rejecting Lewis's account of content-determination undermines an argument to show that our proposal needs modifying before it can accommodate constitutional vagueness. It doesn't follow that our proposal can accommodate that vagueness. This section sketches an account.

Suppose that hair h is clearly part of Tibbles at time t_1, and has fallen out by time t_2. Let T be the lump that constitutes Tibbles at t_1; let T^{-h} be T excluding the matter of h. By t_2, T^{-h} constitutes Tibbles and T is a scattered object. We'll assume for simplicity that Unger's puzzle doesn't arise at t_1 or at t_2: only T bears CON to Tibbles at t_1, and only T^{-h} bears CON to Tibbles at t_2. We'll also assume that h is Tibbles's only borderline part at any time between t_1 and t_2, and that Tibbles undergoes no changes other than those consequent on his loss of h.

When h is a perfectly balanced borderline part of Tibbles, both T and T^{-h} are equally good (and good enough) candidates to constitute Tibbles; they both then bear CON to Tibbles. Lewis's puzzle of borderline constituters thus induces Unger's puzzle of too many best candidates. We want to expand on this to accommodate the Sorites and higher-order vagueness. Our strategy is to mirror the Sharpening theorist's account of typical non-constitutional vagueness.

As h falls out, it gradually becomes less causally integrated with the rest of Tibbles. Underlying this gradually weakening causal connection, is non-gradual variation in CON: it holds from T to Tibbles at t_1, from both T and T^{-h} to Tibbles at some intermediate time(s) t_n, and only from T^{-h} to Tibbles at t_2. The problem was that this imposes sharp boundaries on 'constitutes'. But if we reject the role of

Eligibility in content-determination, then this non-gradual/sharp variation in con needn't translate into sharp boundaries in 'constitutes'. Our use of 'constitutes' is sensitive to the gradually varying causal and spatial relations between h and T^{-h}, not the non-gradual variation in con. Small variations in these respects bring only small variations in fit with our use of 'constitutes': our use of 'constitutes' imposes a non-gradual classification onto this gradual series without privileging any one point in the series over all others. Hence, from the Sharpening theorist's perspective, just the same features that lead to vagueness in 'red', 'old' and 'tall' lead to vagueness in 'constitutes'. Without a role for Eligibility, the cases are alike, and there's no bar to applying the Sharpening View of vagueness. We can therefore accommodate vague constitution without modifying our proposal or the Sharpening View, and without an abundance of cats, provided we reject Lewis's account of content-determination.

4.5.4 Limiting constitutional unclarity

We've got one account of constitutional vagueness in place that modifies neither our proposal nor the Sharpening View. Since that view turns on rejecting Lewis's Eligibility-based account of content-determination, it won't be acceptable to all. So this section develops a different way of extending our proposal to constitutional vagueness.

The view developed here accepts the arguments for the incompatibility of our solution to Unger's puzzle with the Sharpening View of constitutional vagueness. This is taken to show that constitution cannot be vague.[24] The task is to explain away the appearance of constitutional vagueness. This is the goal of §4.5.4.1. We'll do so by adopting an epistemicist strategy. Two further challenges then arise:

> If there cannot be borderline cases to the borderline cases, why allow borderline cases of constitution at all? Why not have sharp boundaries at the first level if we're going to have them anywhere (especially somewhere so close as the second level)?

[24] A notational variant draws the alternative conclusion that vagueness is not a uniform phenomenon.

> Why doesn't this account generalise to all vagueness, and thereby undermine the Sharpening View?

These are addressed in §§4.5.4.2–4.5.4.3.

4.5.4.1 First challenge: the Sorites

Consider a Sorites series on the constitution of Tibbles by T, originating at (i) a time t_1 when h was clearly part of Tibbles, who was then constituted by T, and terminating with (ii) a time t_n when h clearly wasn't part of Tibbles, who was then constituted by T^{-h}. The first challenge is to explain why the following is intuitively plausible, although provably false, and to do so despite the non-gradualness of CON:

(10) $\forall t_i(T \text{ constitutes Tibbles at } t_i \rightarrow T \text{ constitutes Tibbles at } t_{i+1})$

We can meet this challenge by co-opting an epistemicist strategy. Our exposition of this strategy will ignore the unclarity in constitution that arises when both T and T^{-h} bear CON to Tibbles; in other words, we 'll assume that CON holds first from T to Tibbles, and then from T^{-h} to Tibbles, and never from both to Tibbles. It follows that Unger's puzzle never arises. Nothing of substance turns on this, but it simplifies exposition and makes our task harder by providing a clear counterexample to (10). We'll also assume that everything is named.

Our strategy is as follows. First, we'll explain why no instance of the following is knowable (to beings like ourselves), despite one of them being clearly true:

(11) $T \text{ constitutes Tibbles at } t_\alpha \wedge \neg T \text{ constitutes Tibbles at } t_{\alpha+1}$

Then we'll postulate a (mistaken) slip from the unknowability of these instances to their falsity, and provide an explanation of why we make this mistake. Now, if every instance of (11) is false, then so is:

(12) $\exists t_i(T \text{ constitutes Tibbles at } t_i \wedge \neg T \text{ constitutes Tibbles at } t_{i+1})$

And if that is false, then (10) is true. The result is an (invalid but natural) argument from the unknowability of instances of (11) to the truth of the Sorites principle (10). We'll explain the attraction of that Sorites principle by attributing this kind of reasoning to ordinary speakers. Let's turn to the details.

Why is no instance of (11) knowable (to beings like ourselves), despite one being clearly true? Suppose we know all the facts about causal integration, spatial separation and the like that concern T and h. Suppose also that we know exactly what kinds of change Tibbles survives (under a one-level mode of presentation). These are all the facts relevant to CON. So unless the extension of CON is knowable on the basis of these facts, it isn't knowable at all (to beings like ourselves). But the extension of CON isn't knowable on that basis unless we know how the basis bears on CON. Since we don't know that, and its not clear how we might find it out, any means of inferring the extension of CON from these facts would be no better than a guess, even if it gave accurate results; and even an accurate guess doesn't yield knowledge. So we can't know the extension of CON, and hence can't know any instance of (11).

The next task is to explain the slip from the unknowability of any instance of (11) to the falsity of each. Note first that we either do or could in principle know all the facts relevant to the truth of instances of (11). Yet no amount of investigation into those facts would reveal which instance was true. Since this unknowability isn't the result of our own limitations, there must be no truth there to know. So each instance of (11) is false; so (12) is false; so the Sorites principle (10) is true.

Attributing this kind of reasoning to ordinary speakers allows us to explain the slip from the unknowability of each instance of (11) to the truth of (10), despite its falsity. The flaw in the argument is that there's a kind of fact relevant to the extension of CON that we don't know: how what we do know bears on CON. This approach to the Sorites thus attributes forgetfulness or ignorance about the existence of these facts to ordinary speakers.[25] By doing so, we can explain the attraction of constitutional Sorites principles without appeal to (what the Sharpening theorist regards as genuine) vagueness in constitution, and without modifying our response to Unger. The following two sections elaborate this view by responding to the two challenges immediately preceding this section.

[25] One candidate explanation for this ignorance might attribute a form of microphysicalism to ordinary speakers: all relations between macroscopic and microscopic entities are revealed by microphysical descriptions.

4.5.4.2 Second challenge: borderline precision

§4.5.1 argued from our thesis of multiple-constitution to borderline cases of constitution: the many best candidates that all bear con to Tibbles are borderline cases of cat-constituters. The present view posits a sharp boundary between the clear non-cat-constituters and the borderline cat-constituters. Why is this preferable to a sharp boundary between the constituters and the non-constituters? The answer is that it isn't.

On the present view, there is a significant sense in which the constituter/non-constituter distinction is non-vague: there are no borderline cases of con. Our argument from multiple-constitution (i.e. from con being many-one) to borderline constitution didn't appeal to vagueness or the imposition of an absolute classification onto a gradual transition. It appealed instead to the reconciliation of conflicting semantic pressures when determining the extension of 'constitutes': we speak as if constitution were one-one; con isn't one-one; yet con is the best candidate to occupy the constitution-role. If it weren't for these conflicting pressures, we could dispense with unclarity in constitution: the lumps con-related to Tibbles would clearly constitute him, and everything else would clearly fail to.

The moral is that the present approach to (apparent) constitutional vagueness doesn't treat first- and second-order vagueness differently. It denies the existence of both, and hence posits borderline cases in response to neither. It does treat first- and second-order unclarity differently, but that's because the argument for the former doesn't extend to an argument for the latter: the unclarity in constitution that arises from Unger's puzzle is not a form of vagueness.

4.5.4.3 Third challenge: generalisation to other cases

The Sharpening theorist who endorses this approach must explain why it doesn't extend to all other forms of vagueness: why adopt the Sharpening View at all, if an alternative is adequate? This challenge can be met by pointing to a disanalogy with typical cases of vagueness, like 'red'.

Beneath *h*'s gradual working loose lies a non-gradual and highly Eligible distinction between the times when *T* bears con to Tibbles, and those when it doesn't.

CON serves as a "reference magnet" for ordinary constitutional vocabulary', imposing precision on 'constitutes' despite our messy use of language. In this respect, the Sharpening theorist should claim, 'constitutes' is unlike 'red', 'tall' and 'young'. Gradual variation in shade, height and age do not mask any highly Eligible non-gradual distinction; nothing plays the role of CON in imposing a sharp boundary on our messy use of 'red', 'tall' and 'young'. If this is correct, then the present strategy of explaining away the appearance of constitutional vagueness doesn't extend to those cases.

4.5.5 Gradual constitution

We've seen that we can accommodate constitutional vagueness by rejecting Lewis's Eligibility-based account of content-determination (§4.5.3). We've also seen that we can explain away the appearance of constitutional vagueness if we retain that account of content-determination (§4.5.4). This section presents our third and final method for accommodating constitutional vagueness. It generalises our solution to Unger's puzzle by allowing CON to be gradual. Vagueness in 'constitutes' can then be treated in just the same way as for any other form of vagueness: the result of our imposing an absolute classification onto a gradual transition. §4.5.5.1 introduces the proposal. §4.5.5.2 turns to an objection. §4.5.5.3 closes by examining the proposal's interaction with relativised instantiation.

4.5.5.1 The proposal

According to the Sharpening theorist, vagueness results from our imposition of absolute classifications onto a gradual world. The problem with accommodating constitutional vagueness within our solution to Unger's puzzle was that CON is non-gradual. Were CON gradual, there would be no problem. So why not let CON be gradual, and thereby eliminate the problem?

What exactly would it be for CON to be gradual? Consider the gradual transition of shades from orange to red on a colour chart. This gradualness consists in two things. One is the instantiation of many different determinates of the determinable *colour*. The other is an ordering on these determinates.[26] We can apply this to CON, as follows.

There are many determinate CON-relations, each belonging to the same determinable. There is also an ordering on these CON-determinates. For simplicity we'll assume that this ordering is total and dense. Nothing of substance turns on this, but it allows us to write as if a single CON-relation held to a degree d, where d is a real number in the interval $[0,1]$; larger numbers represent stronger constitutional-connections (greater elements in the ordering on CON-determinates).

Our two-level proposal may seem to face a problem with gradual constitution; for that view treats constitution as a function from matter to ordinary objects, and functional-application is non-gradual. However, for each function f there is a functional relation R_f such that:

$f(x) = y$ iff $R_f(x,y)$.

An ontology of functions is thus eliminable in favour of functional relations and function signs governed by the rule:

$\ulcorner f(\alpha)\urcorner$ denotes the unique object y such that $R_f(x,y)$, where x is the referent of α.

Then we can rewrite two-level criteria thus:

$$\forall x \forall y \exists z \exists z'[(R_f(x,z) \land R_f(y,z') \land z = z') \leftrightarrow R(x,y)]$$

In our two-level proposal, R_f is CON (as restricted to an ordinary kind K). That proposal can therefore allow gradual constitution.

Once CON is gradual, the Sharpening View can be applied. Neither our use 'constitutes' nor CON itself privileges some unique degree of CON over all others. This gives rise to many intended interpretations of 'constitutes'. Sorites principles are attractive because they seem to report the absence of relevant differences between successive cases in a Sorites series. Higher-order vagueness arises because the metasemantic facts that determine the intended interpretations of 'intended interpretation of 'constitutes' ' are gradual.

[26] On typical colour charts, the left-right ordering of exemplars of shades matches the ordering on shades themselves. The gradual transition amongst shades is mirrored in their layout on the chart. The ordering on shades is also multi-dimensional, but we'll ignore this complication for simplicity.

4.5.5.2 A limit on higher-order borderline cases?

Although gradualness in CON may allow vagueness in 'constitutes', this proposal may appear to impose a limit on the extent of higher-order vagueness. The problem is that the distinction between standing in CON to no degree and doing so to some degree is non-gradual. This may seem to limit metasemantic gradualness, and hence also higher-order vagueness.

To illustrate the problem, let R, R^* be candidate extensions for 'constitutes' that differ only as follows:

> Let x be a lump that bears CON to Tibbles to degree 0. Let y be a lump that bears CON to Tibbles to some degree only just greater than 0. R and R^* both hold from y to Tibbles, but only R^* holds from x to Tibbles.

The question is: does this suffice to make R^* significantly less Eligible than R? If so, then interpretations that assign R^* to 'constitutes' will fit the meaning-determining facts significantly less well than those that assign R to 'constitutes'. This limits metasemantic gradualness: no series of interpretations, each of which fits the meaning-determining facts only slightly less well than its predecessor, connects interpretations of the following kinds:

> The s-extension of 'constitutes Tibbles' includes something that bears CON to Tibbles to degree 0.

> The s-extension of 'constitutes Tibbles' includes something that bears CON to Tibbles to some degree only just greater than 0.

The result is that objects that don't bear CON to Tibbles at all will be absolutely clearly non-constituters of Tibbles, and no borderline cases will separate them from everything else.

There are two kinds of response we might take. The first denies that this limit on higher-order vagueness is an objectionable limit. The second denies that the argument for this limit is sound. We take them in turn.

Is this limit on higher-order vagueness objectionable? One reason to think not appeals to a version of the view in §4.5.4: this limit isn't the result of our use of language, but is imposed by the underlying facts about con. In particular, it's imposed by the non-gradual distinction between standing in con to some degree, and standing in con to no degree. In order to be objectionable, a limit on higher-order vagueness would have to result from our use of language. Since this one doesn't, it isn't objectionable.

The second kind of response denies that the difference between R and R^* suffices for a significant difference in their Eligibility to be interpretations of 'constitutes'. It's hard to argue either way, given how little is known about Eligibility. It's even more difficult if the Eligibility-ordering is identified with the naturalness-ordering. One reason is that the naturalness-ordering is defined using perfect naturalness, which is supposed to be primitive. Another reason is that it's unclear how perfect naturalness determines the naturalness-ordering.[27] But still, the difference between being con-related to Tibbles to some arbitrarily small degree and not being con-related to Tibbles at all doesn't look like a very significant objective difference. Consider the change from one state to the other. No gradual shift may accompany this change, but it doesn't follow that it's a very significant change: it may not correspond to any major variation in the intrinsic nature of the object in question. It needn't even be a greater change than a change in the degree to which something bears con to Tibbles, if the ordering on con-determinates isn't dense.

In light of these considerations, both the following are doubtful: (a) allowing con to be gradual limits the extent of higher-order vagueness; (b) any limits on higher-order vagueness resulting from our gradual account of con are objectionable limits.

4.5.5.3 Relativised property-possession

§4.4 defended the following view: if Tibbles inherits a property ϕ from a constituter l, then he doesn't have ϕ *simpliciter*, but only relative to l. How does this interact with gradual constitution? There seem to be two suggestions:

Tibbles has ϕ relative to l iff l bears con to Tibbles to some degree greater

[27] Lewis (1986b, p.61) suggests the following: ϕ is more natural than ψ iff ϕ can be reached by a shorter and (or?) less-complicated chain of definability from the perfect naturals than can ψ. But what is the objective standard for complexity and length of a definition?

than n and has ϕ.

Tibbles has ϕ relative to l to degree d iff l both bears CON to Tibbles to degree d and has ϕ.

We should reject the first. On that view, if l bears CON to Tibbles to less than degree n, then there's no sense in which Tibbles has the same mass, shape, location and so on as l. It's mysterious how l could then count as even remotely constitutionally connected to Tibbles. So let's consider the second view.

On our preferred account of relativised possession (§4.4.4.4), property, lump and object all enter into a single structure; this structure isn't analysable into a relation's obtaining amongst its constituents. When combined with gradual constitution, this yields as many different varieties of these structures as there are CON-determinates. There seem to be two problems with this. The first is that it brings a massive increase in our theory's primitive ideology. The second is that it's unclear what these structures all have in common: why do they all count as relativised-possession-structures? Were they analysable using an instantiation relation I, then we could appeal to different determinates of the determinable I. But that analysis is just what our preferred view of relativised possession denies. The defender of gradual constitution should therefore prefer the alternative account of relativised possession (§4.4.4.3). On that view, relativised possession is analysable in terms of constitution and the properties of matter. This allows us to take the r.h.s. of the second biconditional above as an analysis of the left, and hence of relativised possession to a degree without any ideological cost. The defender of gradual constitution can then accommodate vague constitution, higher-order borderline cases of constitution, and degrees of relativised possession simply by appeal to the gradualness of CON. This completes our third account of vague constitution. The Sharpening theorist who is prepared to allow gradual constitution can accommodate the vagueness of constitution without an abundance of cats.

4.6 Conclusion

This chapter presented several solutions to the Problem of the Many. Each develops the thesis that change is explanatorily prior to constitution. §4.2 presented one-level and two-level identity criteria as ways of developing this view. The key difference between these two views lies in what kinds of change they claim take priority over constitution: changes in the persisting object itself, or changes in its matter. Although the two views aren't in direct competition, §4.3.5 argued that the one-level view is preferable, without ruling out the two-level view entirely. The choice between these views also doesn't affect our solution to Unger's puzzle: Tibbles is constituted by each of the best candidates on his mat. Both the one-level view and the two-level allow us to mount direct arguments for this claim. This shows that, unlike Lowe and Johnston's proposals, ours is not merely an arbitrary collection of theses designed to invalidate the arguments for many cats. §4.4 finished the exposition of our solution to Unger's puzzle by relativising Tibbles's inherited properties, like mass, colour and location, to the matter from which he inherits them.

§4.5 closed with a discussion of unclarity and vagueness. We argued from our proposal to unclarity in constitution, mereology and inherited properties. These arguments exploit a mismatch between linguistic structure and the structure of the underlying facts in order to locate many equally suitable interpretations of the vocabulary in question. We then argued that this unclarity isn't genuine vagueness, but another form of linguistic unclarity: Unger's puzzle is not directly a puzzle of vagueness. We closed with three ways of extending our solution to Unger's puzzle to constitutional vagueness. The first relied on a rejection of Lewis's Eligibility-based conception of content-determination. The second rejected constitutional vagueness and attempted to explain away its appearance without undermining the Sharpening View. The third modified our solution to Unger's puzzle by allowing constitution to be a gradual matter. Each of these views has its own costs and benefits, which there isn't space here to evaluate properly. Whichever of these views we prefer, the result is a unified solution to both Unger's and Lewis's Problems of the Many on which there is only ever one cat on Tibbles's mat, and a conception of ordinary objects from which this solution emerges naturally.

Bibliography

Barnett, David (2008). "Indeterminacy and incomplete definitions". *The Journal of Philosophy*, 55: 167–91.

Blackburn, Simon and Simmons, Keith, eds. (1999). *Truth*. OUP.

Boolos, George (1984). "To be is to be the value of a variable (or some values of some variables)". *The Journal of Philosophy*, 81: 430–49.

Chisholm, Roderick M. (1976). *Person and Object: A Metaphysical Study*. George Allen & Unwin.

Cook, Roy T. (2002). "Vagueness and mathematical precision". *Mind*, 111: 225–247.

Dever, Josh (2009). "The disunity of truth". In Robert J. Stainton and Christopher viger, eds., *Compositionality, Context and Semantic Values: Essays in Honour of Ernie Lepore*, chapter 7. Springer Netherlands.

Dorr, Cian (2003). "Vagueness without ignorance". *Philosophical Perspectives 17: Language and Philosophical Linguistics*, 83–113.

Edgington, Dorothy (1997). "Vagueness by degrees". In Rosanna Keefe and Peter Smith, eds., *Vagueness: a reader*, chapter 16. MIT Press.

—— (2010). "Possible knowledge of unknown truth". *Synthese*, 173: 41–52.

Eklund, Matti (2008). "Deconstructing ontological vagueness". *Canadian Journal of Philosophy*, 38: 117–40.

Etchemendy, John (1990). *The Concept of Logical Consequence*. Cambridge Mass., London England: Harvard University Press.

Evans, Gareth (1978). "Can there be vague objects?" *Analysis*, 38: 208.

Field, Hartry (1974). "Quine and the correspondence theory". *The Philosophical Review*, 83: 200–28.

—— (2008). *Saving Truth From Paradox*. Oxford: OUP.

Fine, Kit (1975). "Vagueness, truth and logic". *Synthese 30*, 265–300.

—— (1999). "Things and their parts". *Midwest Studies in Philosophy*, 23: 61–74.

—— (2000). "A counter-example to Locke's Thesis". *The Monist*, 83: 357–61.

—— (2003). "The non-identity of a material thing and its matter". *Mind*, 112: 195–234.

—— (2008). "Coincidence and form". *The Aristotelian Society: Supplementary Volume*, 82: 101–18.

Geach, Peter (1980). *Reference and Generality*. 3rd. edition. Ithaca, NY: Cornell University Press.

Hawley, Katherine (2001). *How Things Persist*. Oxford, New York: OUP.

Hawthorne, John (2006a). "Epistemicism and semantic plasticity". *Oxford Studies in Metaphysics*, 2: 289–324. Reprinted in Hawthorne (2006b, ch.9).

—— (2006b). *Metaphysical Essays*. Clarendon Press, Oxford: OUP.

—— (2006c). "Plenitude, convention, and ontology". In Hawthorne (2006b, ch.3).

—— (2006d). "Three-dimensionalism". In Hawthorne (2006b, ch.5).

Hazen, Allen (1979). "Counterpart-theoretic semantics for modal logic". *Journal of Philosophy*, 76: 319–38.

Hirsch, Eli (1982). *The Concept of Identity*. OUP.

Horsten, Leon (2010). "Impredicative Identity Criteria". *Philosophy and Phenomenological Research*, 80: 411–39.

Hudson, Hud (2001). *A Materialist Metaphysics of the Human Person*. Ithaca, NY: Cornell University Press.

Hughes, Christopher (1986). "Is a thing just the sum of its parts?" *Proceedings of the Aristotelian Society*, 86: 213–33. Reprinted in Noonan (1993, pp.79–99).

Hyde, Dominic (1997). "From heaps and gaps to heaps of gluts". *Mind*, 108: 641–80.

Johnston, Mark (1992). "Constitution is not identity". *Mind*, 101: 89–105. Reprinted in Rea (1997, pp.44–62).

Jones, Nicholas K. (forthcoming). "Williams on supervaluationism and logical revisionism". *Journal of Philosophy*.

Keefe, Rosanna (2000). *Theories of Vagueness*. Cambridge: Cambridge University Press.

Keefe, Rosanna and Smith, Peter, eds. (1997). *Vagueness: A Reader*. MIT Press.

Kripke, Saul A. (1982). *Wittgenstein on Rules and Private Language*. Oxford: Blackwells.

LePore, Ernest and Fodor, Jerry (1996). "What cannot be evaluated cannot be evaluated, and it cannot be supervalued either". *Journal of Philosophy*, 93: 513–36.

Lewis, David (1968). "Counterpart theory and quantified modal logic". *The Journal of Philosophy*, 65: 113–26. Reprinted in Lewis (1983b, ch.3).

—— (1969). *Convention: A Philosophical Study*. Cambridge, Mass.: Harvard University Press.

—— (1970a). "General Semantics". *Synthese*, 22: 18–67.

—— (1970b). "How to define theoretical terms". *The Journal of Philosophy*, 67: 427–46. Reprinted in Lewis (1983b, ch.6).

—— (1971). "Counterparts of persons and their bodies". *The Journal of Philosophy*, 68: 203–11. Reprinted in Lewis (1983b, ch.4).

—— (1976). "Survival and identity". In Amelie O. Rorty, ed., *The Identities of Persons*, 17–40. University of California Press. Reprinted in Lewis (1983b, ch.5).

—— (1983a). "New work for a theory of universals". *The Australasian Journal of Philosophy*, 61: 343–77. Reprinted in Lewis (1999, pp.8–55).

—— (1983b). *Philosophical Papers: Volume 1*. New York, Oxford: OUP.

—— (1984). "Putnam's Paradox". *The Australasian Journal of Philosophy*, 62: 221–36. Reprinted in Lewis (1999, pp.56–77).

—— (1986a). *Counterfactuals*. revised edition. Cambridge MA: Harvard University Press.

—— (1986b). *On the Plurality of Worlds*. Malden MA, Carlton Victoria, Oxford: Blackwell Publishing.

—— (1988a). "Rearrangement of particles: reply to Lowe". *Analysis*, 48: 65–72. Reprinted in Lewis (1999, ch.11).

—— (1988b). "Vague identity: Evans misunderstood". *Analysis*, 48: 128–30. Reprinted in Keefe and Smith (1997, ch.18).

—— (1991). *Parts of Classes*. Oxford: Basil Blackwell.

—— (1993a). "Many, but almost one". In *Ontology, Causality and Mind: Essays on the Philosophy of D.M. Armstrong*, 23–37. Cambridge: Cambridge University Press. Reprinted in Lewis (1999, ch.9).

—— (1993b). "Mathematics is megethology". *Philosophia Mathematica*, 1: 3–23.

—— (1999). *Papers in Metaphysics and Epistemology*. Cambridge University Press.

—— (2002). "Tensing the copula". *Mind*, 111: 1–13.

—— (2009). "Ramseyan humility". In David Braddon-Mitchell & Robert Nola, ed., *Conceptual Analysis and Philosophical Naturalism*, chapter 9. MIT Press.

Loux, Michael J. and Zimmerman, Dean (2003). *Oxford Handbook of Metaphysics*. Oxford, New York: OUP.

Lowe, E. J. (1982a). "On being a cat". *Analysis*, 42: 174–77.

—— (1982b). "Reply to Geach". *Analysis*, 42: 31.

—— (1982c). "The paradox of 1,001 cats". *Analysis*, 42: 27–30.

—— (1989). "What is a criterion of identiy?" *The Philosophical Quarterly*, 39: 1–21.

—— (1991). "One-level versus two-level identity criteria". *Analysis*, 51: 192–4.

—— (1995). "The problem of the many and the vagueness of constitution". *Analysis*, 55: 179–82.

MacBride, Fraser (2001). "Four new ways to change your shape". *Australasian Journal of Philosophy*, 79: 81–9.

—— (2005a). "Review of Herbert Hochberg, *Introducing Analytic Philosophy: Its Sense and its Nonsense 1879–2002*". *Dialectica*, 59: 81–87.

—— (2005b). "The particular-universal distinction: a dogma of metaphysics?" *Mind*, 114: 565–614.

Markosian, Ned (1998). "Brutal composition". *Philsophical Studies*, 92: 211–249.

McFetridge, Ian (1990). "Logical necessity: some issues". In *Logical Necessity and Other Essays*, number 10 in Aristotelian Society Series, 135–54. London: Aristotelian Society. Edited by John Haldane and Roger Scruton.

McGee, Vann and McLaughlin, Brian (1994). "Distinctions without a difference". *The Southern Journal of Philosophy*, 33: 203–261.

—— (1998). "Review of Williamson (1994)". *Linguistics and Philosophy*, 21: 221–35.

—— (2000). "The lessons of the Many". *Philosophical Topics*, 28: 129–51.

—— (2004). "Logical commitment and semantic indeterminacy: a reply to Williamson". *Linguistics and Philosophy*, 27: 122–36.

McKinnon, Neil (2002). "Supervaluations and the problem of the many". *The Philosophical Quarterly*, 52: 320–39.

Merricks, Trenton (2003). *Objects and Persons*. Clarendon Press: OUP.

Moore, Jospeh (1999). "Propositions without identity". *Nous*, 33: 1–29.

Noonan, Harold, ed. (1993). *Identity*. Dartmouth Publishing Company.

Putnam, Hilary (1980). "Models and reality". *Journal of Symbolic Logic*, 45: 464–82.

Quine, W. V. O. (1976). "Worlds away". *The Journal os Philosophy*, 73: 859–65.

—— (1981a). *Theories and Things*. Harvard University Press.

—— (1981b). "What price bivalence?" *Journal of Philosophy*. Reprinted with additions in Quine (1981a, ch.3).

Rea, Michael C., ed. (1997). *Material Constitution: A Reader*. Lanham, Boulder, New York, Oxford: Rowman & Lanham Publishers inc.

Rumfitt, Ian (1993). "Content and context: the paratactic theory revisited and revised". *Mind*, 102: 429–54.

—— (2002). "Concepts and counting". *Proceedings of the Aristotelian Society*, 102: 41–68.

Sainsbury, R. M. (1990). *Concepts without boundaries*. King's College London Depatment of Philosophy. Reprinted in Keefe and Smith (1997, pp.251–64).

Schiffer, Stephen (1998). "Two issues of vagueness". *The Monist*, 81: pp.193–214.

—— (2000). "Replies". *Philosophical Issues*, 10: 321–43.

Shapiro, Stewart (1998). "Logical consequence: models and modality". In Matthias Schirn, ed., *The Philosophy of Mathematics Today*, 131–56. Clarendon Press: OUP.

—— (2006). *Vagueness in Context*. Clarendon Press: OUP.

Sharvy, Richard (1980). "A more general theory of definite descriptions". *The Philosophical Review*, 89: 607–24.

Sider, Theodore (2001a). *Four-Dimensionalism: An Ontology of Persistence and Time*. Oxford, New York: OUP.

—— (2001b). “Maximality and intrinsic properties”. *Philosophy and Phenomenological Research*, 63: 357–64.

—— (2003). “Maximality and microphysical supervenience”. *Philosophy and Phenomenological Research*, 66: 139–49.

Simons, Peter (1987). *Parts: A Study in Ontology*. Clarendon Press: OUP.

—— (2000). “Continuants and occurrents”. *Proceedings of the Aristotelian Society: Supplementary Volume*, 74: 59–75.

—— (2008). “The thread of persistence”. In Christian Kanzian, ed., *Persistence*, 165–84. Ontos Verlag.

Sorensen, Roy (2000). “Direct reference and vague identity”. *Philosophical Topics*, 28: 175–94.

—— (2010). “Borderline hermaphrodites: higher-order vagueness by example”. *Mind*, 119: 393–408.

Stalnaker, Robert C. (1986). “Counterparts and Identity”. *Midwest Studies in Philosophy: Studies in Essentialism*, 11: 121–40. Reprinted in Stalnaker (2003, ch.6).

—— (1994). “The interaction of quantification with modality and identity”. In N. Asher W. Sinnott Armostrong, D. Raffman, ed., *Modality, Morality and Belief: Essays in honor of Ruth Barcan Marcus*. Cambridge: Cambridge University Press. Reprinted in Stalnaker (2003, ch.8).

—— (2003). *Ways a World Might Be: Metaphysical and Anti-Metaphysical Essays*. Clarendon Press: OUP.

Tarski, Alfred (1936). “Über den begriff der logischen folgerung”. *Actes du Congrès International de Philosophie Scientifique*, 7: 1–11. Translated as “On the concept of logical consequence” in Tarski (1956).

—— (1956). *Logic, Semantic, Metamathematics*. Oxford: Clarendon Press.

Tye, Michael (1996). "Fuzzy realism and the problem of the many". *Philosophical Studies*, 81: 215–25.

Unger, Peter K. (1979). "Why there are no people". *Midwest Studies in Philosophy*, IV: 177–222. Reprinted in Unger (2006b, pp.53-111).

—— (1980). "The problem of the many". *Midwest Studies in Philosophy*, 5: 411–67. Reprinted in Unger (2006b, pp.113–83).

—— (2006a). *All the Power in the World*. Oxford, New York: OUP.

—— (2006b). *Philosophical Papers: Volume 2*. Oxford, New York: OUP.

van Inwagen, Peter (1990). *Material Beings*. Ithaca, London: Cornell University Press.

Weatherson, Brian (2003). "Many many problems". *The Philosophical Quarterly*, 53: 481–501.

—— (2009). "The Problem of the Many". In Edward N. Zalta, ed., *The Stanford Encyclopedia of Philosophy*, summer 2009 edition. Accessed October 2009. URL `http://plato.stanford.edu/archives/sum2009/entries/problem-of-many/`

Wiggins, David (1968). "On being in the same place at the same time". *The Philosophical Review*, 77: 90–95. Reprinted in Rea (1997, ch.1).

—— (1980). *Sameness and Substance*. Oxford: Basil Blackwell.

—— (2001). *Sameness and Substance Renewed*. Cambridge Universit Press.

Williams, J. R. G. (2006). "An argument for the many". *Proceedings of the Aristotelian Society*, 106: 411–19.

—— (2008). "Supervaluationism and logical revisionism". *Journal of Philosophy*, 55: 192–212.

Williamson, Timothy (1990). *Identity and Discrimination*. Basil Blackwell.

—— (1994). *Vagueness*. London, New York: Routledge.

—— (1997). "Imagination, stipulation and vagueness". *Philosophical Issues*, 8: 215–28.

—— (1999). "On the structure of higher-order vagueness". *Mind*, 108: 127–43.

—— (2003). "Vagueness in Reality". In Loux & Zimmerman Loux and Zimmerman (2003), 690–715.

—— (2004). "Reply to McGee and McLaughlin". *Linguistics and Philosophy*, 27: 113–22.

—— (2007). *The Philosophy of Philosophy*. Blackwell Publishing.

Wright, Crispin (1976). "Language-mastery and the sorites paradox". In G Evans and J McDowell, eds., *Truth and Meaning: Essays in Semantics*, 223–47. Oxford, Clarendon Press: OUP. Reprinted in Keefe and Smith (1997, ch.10).

—— (1999). "Truth: a traditional debate reviewed". *Canadian Journal of Philosophy*, 24. Reprinted in Blackburn and Simmons (1999, ch.13).

Zimmerman, Dean W. (1995). "Theories of Masses and Problems of Constitution". *The Philosophical Review*, 104: 53–110.