Vagueness and the Philosophy of Perception

Michael Ryan Perkins

m.ryan.perkins@gmail.com

* This dissertation was supervised by John Hawthorne, to whom the writer is eternally grateful. Thanks of equal measure are due to Jeremy Goodman, for hundreds of hours of discussion, and to Tim Bayne, Kai Spiekermann, Matthew Soteriou, Bill Brewer, and Guy Longworth.

ABSTRACT

This dissertation explores several illuminating points of intersection between the philosophy of mind and the philosophy of vagueness. Among other things, I argue:

- That a popular recent account of perceptual phenomenology (representationalism) conflicts with our best theory of vagueness (supervaluationism);
- That there are no vague properties; and
- That strong versions of dualism are unable to accommodate the possibility of borderline consciousness.

Introduction
(*pp*. 4-6)

CHAPTER ONE
***What It's Like*: A Locution Misunderstood**
(*pp*. 7-31)

CHAPTER TWO
Representationalism and the Problem of Vagueness
(*pp*. 32-51)

CHAPTER THREE
Vague Properties
(*pp*. 52-72)

CHAPTER FOUR
Δ-Representationalism
(*pp*. 73-85)

CHAPTER FIVE
Dualism and Borderline Consciousness
(*pp*. 86-107)

Bibliography
(*pp*. 108-115)

Introduction

This dissertation explores several illuminating points of intersection between the philosophy of mind and the philosophy of vagueness.

Several theorists have hinted at the potential for fruitful interaction between the philosophy of mind and the philosophy of vagueness. For example, there is an absorbing literature on the 'phenomenal sorites' (e.g. Graff 2001). The relationship between physicalism and vagueness has also received discussion in certain quarters (e.g. Papineau 1993; Antony 2006). The view that consciousness admits borderline cases is discussed, albeit briefly, by McGinn (1996: 14) and others. On the whole, however, few theorists have approached philosophical debates about perception and mind through the prism of vagueness at any length. This dissertation takes steps towards rectifying the discrepancy.

A brief summary of each chapter is supplied below. (For the moment, ignore CHAPTER ONE.)

- CHAPTER TWO argues that *representationalism* (a popular account of perceptual phenomenology) conflicts with *precise supervaluationism* (our best theory of vagueness). Indeed, given plausible assumptions, representationalism and precise supervaluationism are provably inconsistent. This is the 'problem of vagueness'; it poses a serious threat to representationalism.

The following two chapters explore whether representationalists are capable of providing an adequate response to the problem of vagueness.

- CHAPTER THREE considers a response which embraces the existence of so-called 'vague properties'. The response fails: there are strong reasons to deny that such properties exist.
- CHAPTER FOUR considers an alternative response, which claims that representationalism—though true—is not *fully determinate*. The response faces several challenges. Nevertheless, it merits further investigation.

This concludes our discussion of representationalism.

- CHAPTER FIVE raises two (related) questions. First: does consciousness admit borderline cases? As we shall see, there is good *prima facie* reason to believe that the answer is 'yes'. Second: which theories are compatible with the view that consciousness admits borderline cases? I shall argue that certain versions of dualism cannot accommodate borderline consciousness, though physicalists face no such difficulty.

Much of the dissertation is conducted in an exploratory spirit, with a view to expansion and further investigation in the future.

This dissertation assumes that readers are confident users of first-order logic. We shall devise formal or quasi-formal proofs of numerous claims in CHAPTER TWO and CHAPTER THREE; reasoning in the presence of vagueness is so challenging that it would be irresponsible to do otherwise.

In addition, this dissertation assumes that properties exist, and that properties are plenitudinous (as opposed to sparse).[1] We shall regularly make use of the first-order predicate '*x* instantiates *y*', without prejudice to its ultimate analysis. (Perhaps properties are universals and '*x* instantiates *y*' is a primitive predicate; perhaps properties are sets and '*x* instantiates *y*' is equivalent to '*x* is a member of *y*'; perhaps properties are sets of tropes and '*x* instantiates *y*' is definable in terms of set-membership and parthood.)

Some philosophers (e.g. Williamson 2004) prefer to regiment talk of properties in higher-order languages. Such theorists decline to accept the first-order sentence: $\exists x \exists y$(*x* instantiates *y*), and instead accept a surrogate sentence of higher-order logic: $\exists F \exists x(Fx)$. Every important claim in this dissertation can easily be translated into higher-order logic, without loss of substantive content. However, higher-order languages generate irrelevant complications that are best ignored here; we shall stick to first-order logic.

[1] For a careful account of what it means to say that properties are 'plenitudinous', see Lewis (1983). CHAPTER THREE discusses the metaphysics of properties in greater detail.

In the chapter summary provided above, we ignored CHAPTER ONE. Why?

The philosophy of mind is filled with talk of 'phenomenal consciousness', 'phenomenal character', 'phenomenal properties', and so forth. This dissertation is no exception. How should such terms be understood? *That* is the central question discussed in CHAPTER ONE. It is very common to define such terms by appeal to the locution: 'what it's like to ϕ'. I shall argue that this common approach faces a number of great difficulties, and articulate an alternative definition.

CHAPTER ONE contains no discussion of the relationship between the philosophy of vagueness and the philosophy of mind. On the other hand, CHAPTER ONE is far from anomalous: it lays the foundation for the entire dissertation, by providing an intelligible account of the central theoretical terms employed throughout.

What It's Like: A Locution Misunderstood

CHAPTER ONE

§1 Four Puzzles

Following the publication of Thomas Nagel's paper *What Is It Like to Be a Bat?* (1974), virtually every philosophical discussion of phenomenal consciousness has come to place considerable theoretical weight on a family of related locutions: 'what it's like to ϕ', 'what it's like for *A* to ϕ', 'there is something it's like to be *A*', *&c*. There seems to be a rare consensus among philosophers of mind: that *WIL*-talk is an unproblematic tool for theorizing about the metaphysics of consciousness.[2]

I shall present a series of closely related puzzles which challenge the plausibility of this consensus. Collectively, the puzzles suggest that *WIL*-talk is much less philosophically useful than many theorists believe. Of course, the puzzles developed below give rise to a plethora of questions and objections. For structural reasons, it will prove fruitful to begin by briefly considering all the puzzles together, delving more deeply thereafter.

FIRST PUZZLE

Is there life after death? Needless to say, the question is not exactly new. Accordingly, it is disappointing that Anselm and other eminent theologians appear to have missed a very straightforward argument for the afterlife. The argument relies on a widely accepted schema (cf. Nagel 1974):

Schema 1

If there is something that it's like to ϕ, then ϕ-ing requires consciousness.

[2] Henceforth, 'consciousness' is short for 'phenomenal consciousness'; 'access consciousness' (Block 1995) is largely ignored in this dissertation.

(ϕ-ing requires consciousness iff every subject who ϕs is conscious.) We argue for the afterlife as follows. Suppose—for *reductio*—that there is no afterlife. Presumably, then, Avril answers Hal's question correctly:

Hal What is it like to be dead?
Avril It's like being in a dreamless sleep.

Of course, proponents of the afterlife will dispute Avril's answer—but it is hard to see how opponents of the afterlife can have any cause for concern. Indeed, Avril's answer to Hal's question is fairly commonplace.[3] Assuming that Avril's answer is correct, it plainly follows that there is *something* that it's like to be dead: it's like being in a dreamless sleep. By schema 1, all dead people are conscious. This contradicts our initial supposition that there is no life after death.[4] Conclusion: the afterlife exists.

Of course, this argument is presented in a facetious spirit. Even theists will be reluctant to attain afterlife on the cheap. Rather, the argument above appears to show that something is deeply wrong with schema 1, in spite of its enduring popularity.

Schema 1 does not only entail that there is life after death—it also entails *contradictions*. Imagine a super-anaesthetic which causes 'super-sedation': sedation to the point of complete unconsciousness (no dreaming, no thinking). Hal is about to receive the anaesthetic; Avril is his doctor.

Hal What is it like to be super-sedated?
Avril It's like being in a dreamless sleep.

Assuming that Avril's answer is correct, it follows that there is something that it's like to be super-sedated. By schema 1, all super-sedated subjects are conscious. If it is possible to be super-sedated, we have a contradiction. What has gone wrong?

---

[3] Type 'like being in a dreamless sleep' into *google.com*: you will find thousands of websites which proclaim that 'being dead is like being in a dreamless sleep', or some nearby variant.

[4] Perhaps radical panpsychists can agree that dead people are conscious, yet deny that there is life after death. The resolution to this debate turns on esoteric minutiae concerning the semantics of the word 'afterlife'. For discussion, see my unpublishable paper 'Panpsychism and the Afterlife: A Marriage Made in Heaven?'

SECOND PUZZLE

Analogous problems confront related schemata. Consider:

Schema 2

If *A* is a subject, and there is something that it's like <u>for *A* to ϕ</u>, then *A* is conscious.

Schema 2 employs the locution '…it's like for *A* to ϕ'. Nagel hints at this schema when he writes: 'an organism has conscious mental states if and only if there is something that it is like to *be* that organism… something it is like *for* the organism' (1974: 436). Moreover, certain philosophers make the surprising claim that the complementizer 'for' has a special connection to subjectivity (e.g. Rosenthal 2002: 656). Might schema 2 avoid the problems latent in schema 1?

No. Imagine that Hal wakes up from the world's first month-long dose of super-anaesthetic. He is interviewed by a local reporter.

Interviewer What was it like for you to lie in the hospital for all those weeks?
Hal It was just like being in a dreamless sleep.

There is nothing infelicitous about Hal's answer to the interviewer. Evidently, Hal's answer implies that there is something it was like for Hal to lie in the hospital: it was like being in a dreamless sleep. By schema 2 (together with elementary tense logic), it follows that Hal was conscious while super-sedated. This is a contradiction, by the definition of 'super-sedated'. Schema 1 and schema 2 are equally problematic; the complementizer 'for' does not help.

THIRD PUZZLE

Similar remarks apply to:

Schema 3

If *A* is a subject, and there is something that it's like <u>to be *A*</u>, then *A* is conscious.

Schema 3 employs the locution '…like to be *A*'. Against schema 3, imagine that Hal is super-sedated. Consider the following conversation, conducted among two of Hal's visitors:

Child	I hope Hal isn't in pain.
Parent	No, none at all. Don't worry.
Child	Hmm… I wonder what it's like to be Hal right now.
Parent	Just like being in a completely dreamless sleep, I imagine.

There is nothing infelicitous about this conversation. To the contrary, it is perfectly commonplace. Taken at face value, the conversation implies that there is something it's like to be Hal. (Being Hal is like being in a dreamless sleep.) By schema 3, Hal is conscious: the same undesirable implication carried by schemata 1 and 2.

The moral is clear. The differences between the locutions '…it's like to ϕ', '…it's like to be *A*', and '…it's like for *A* to ϕ' are irrelevant for our purposes. These locutions face structurally identical problems.

FOURTH PUZZLE

Talk of 'phenomenal character' (or 'phenomenal properties' or 'qualia') is commonplace in contemporary philosophy of mind. Representationalism itself, along with its rivals, is commonly formulated in phenomenal-character-theoretic language. What, then, are phenomenal characters?

There is a standard answer. David Chalmers (2004: §2) writes that phenomenal characters 'characterize aspects of what it is like to be a subject'. A very similar definition is provided by Ned Block, who writes that 'the totality of the experiential properties of a state are 'what it is like' to have it' (1995: 230). Michael Tye agrees (2009: 303). *Prima facie*, these philosophers would accept the following gloss on their view:

Schema 4

If *A* is a person and *NP* is what it's like for *A* to ϕ, then *NP* is a phenomenal character.

To obtain an instance of schema 4, '*NP*' must be replaced with a noun phrase.

The first thing to observe about schema 4 (and, indeed, about 'what it's like'-talk more generally) is its extreme syntactic complexity. The formula '*NP* is what it's like to ϕ' contains, among other things, an untensed infinitival clause (to ϕ), a difficult dummy pronoun (what it's like), a difficult indirect question (what it's like to to ϕ), and a difficult usage of 'is' in the context of a pseudocleft (*NP* is what it's like to ϕ). To my mind, it is startling that such a confusing expression, which gives rise to so many obscure and unresolved syntactic and semantic issues, has featured so extensively in the study of consciousness.[5]

In any case, schema 4 faces an immediate difficulty. Consider the following conversation:

Nurse (Preparing to inject Avril with a 16-inch needle.) This might hurt—there's a chance you'll undergo a pretty painful experience. To help me monitor your painkillers, tell me what it's like to receive the injection.

Avril (In pain.) ARGH! It's like being stabbed with a massive knife!

According to Avril, being stabbed with a massive knife is what receiving the injection is like. Equivalently, being stabbed with a massive knife is what it's like to receive the injection. By schema 4, it follows that being stabbed with a knife is a phenomenal character.[6]

This is not a welcome result. Zombies and pillows alike can be stabbed with knives. Whatever theoretical role the notion of phenomenal character is designed to play, presumably unconscious beings, such as zombies and pillows, are not allowed to

[5] On the difficulties raised by indirect questions, see Karttunen (1977) and Hamblen (1973). On the difficulties raised by 'wh'-pseudoclefts, see Partee (1986) and Yablo (1996). For an introduction to the interpretation of infinitival clauses, and their silent constituent 'PRO', see Stanley and Williamson (2000).

[6] The sentence 'being stabbed with a knife is a phenomenal character' may appear grammatically odd. The oddity arises because the expression 'being stabbed with a knife' can occur as part of a predicate (e.g. a present continuous verb in the passive), where it does not function as a singular term, as in: 'Avril is being stabbed with a knife.' This contrasts with sentences like: 'being stabbed with a knife is similar to being stabbed with a fork'. In the latter sentence, 'being stabbed with a knife' is a noun phrase, and functions as a singular term. *That* is its intended interpretation in the text.

instantiate phenomenal characters. Being stabbed with a knife cannot be a phenomenal character. What has gone wrong?

By the way: if you find claims of the form '*NP* is what it's like to ϕ' somewhat unnatural, join the club. They *are* somewhat unnatural. So are other 'pseudo-clefts' of the same form:

Obama is who the 44th President is. (Better: Obama is the 44th President.)
4444 is what the password is. (Better: 4444 is the password.)
London is where Big Ben is. (Better: Big Ben is in London.)

If this is a problem for anyone, it is not a problem for me. It's a problem for the proponents of schema 4, who have made an unnatural locution the centrepiece of their philosophy of mind. If *any* sentences of the form '*NP* is what it's like to ϕ' are true, then 'being stabbed with a knife is what it's like to receive the injection' is surely among them.

…

The four puzzles presented above have a common source. On a natural and commonsensical interpretation, 'what it's like'-talk fundamentally concerns *similarity*. *WIL*-talk aims to identify people, events, and properties that are similar along some salient dimension. (After all, the word 'like' is a constituent of 'what it's like' for a reason.) But if *WIL*-talk merely concerns similarity, it is hard to believe that such talk tracks anything interesting about *consciousness*.

So much by way of introduction. The remainder of this chapter considers the puzzles in greater detail, and explores how we ought to react. In my view, the puzzles reveal that *WIL*-talk is an unhelpful and confusing tool for theorizing about consciousness. It would be best, I submit, if the philosophy of mind were completely divested of *WIL*-talk (§**2**).

Many theorists will view this response with suspicion. Without *WIL*-talk, how can we define important theoretical notions in the philosophy of mind like 'phenomenal character'? §**3** develops and defends an alternative analysis of this notion which does not rely on *WIL*-talk. Instead, the analysis relies on the distinction

between determinate and determinable properties. §**4** defends the analysis from a time-honoured objection. Finally, §**5** develops my favoured formulation of the representationalist theory of perceptual phenomenology, making use of the notions developed in previous sections. The moral is clear: theorizing about consciousness needn't be constrained by the confusing semantics and obscure syntax of *WIL*-talk; we can get by perfectly well without it.

§2 Delving Further

Let us consider the puzzles presented in §**1** more carefully, and explore whether the defender of *WIL*-talk has any adequate response. The puzzles raise exactly same issues, so we shall only discuss the first puzzle in detail.

The first puzzle, readers will recall, targeted schema 1 (= 'if there is something that it's like to ϕ, then ϕ-ing requires consciousness'). We imagined a conversation between soon-to-be-sedated Hal and his doctor:

Hal What is it like to be super-sedated?
Avril It's like being in a dreamless sleep.

It appears to follow that there is something that it's like to be super-sedated. This is a serious problem for schema 1. How might one of its defenders dodge the *reductio*?

One initial line of response is to hold that *even though Avril's answer is correct*, there is nothing it's like to be super-sedated. This strategy, I submit, is a non-starter.

Let me be clear: I do not intend to translate Hal and Avril's conversation into a formal language, define a notion of logical consequence, and prove that some relevant formal analogue of the sentence 'there is something that it's like to be super-sedated' is a logical consequence of Avril's answer. As noted above, the syntax and semantics of *WIL*-talk is very poorly understood. To develop a formal language suited for capturing the logical properties of *WIL*-talk would take us far beyond the purview of this chapter.

I propose a different style of argument. Consider: if a child lacks a name, then the only correct answers to the question 'what is her name?' are negative (e.g. 'she

does not have a name'). If the relevant question has a positive correct answer, then the child has a name. If the USA does not have a 45th President, then the only correct answers to the question 'who is the 45th President of the USA?' are negative (e.g. 'there isn't one'). If the relevant question has a positive correct answer, then someone is the USA's 45th President. If you didn't steal the chocolate, then the only correct answers to the question 'why did you steal the chocolate?' are negative ('I didn't steal it'). If the relevant question has a positive correct answer, then you stole the chocolate for some reason. The same basic theme could be continued indefinitely. Most importantly, if there is nothing that it's like to ϕ, then the only correct answers to the question 'what is it like to ϕ?' are negative (e.g. 'nothing'). If the relevant question has a positive correct answer, then there is something that it's like to ϕ. Furthermore, Avril's answer to Hal's question is plainly not negative. I shall not provide any analysis of what 'negative answers' are, but the notion is clear enough to work with, and Avril's answer does not fall within its purview. If her answer is correct, then there must be something that it's like to be super-sedated.

In any case, we can argue that there is something that it's like to be super-sedated without appealing specifically to Avril and Hal's conversation. We might simply reason as follows:

1. Being super-sedated is like being in a dreamless sleep.
2. Therefore, there is something that is like being super-sedated.
3. Therefore, there is something that it is like to be super-sedated.

(1) is plainly true, (1) plainly entails (2), and (2) plainly entails (3).
(2) and (3) differ only in irrelevant syntactic respects. In particular, (3) contains a dummy pronoun ('it') and an infinitive ('to be super-sedated'), whereas (2) contains no dummy pronoun and a gerund ('being dead') in place of the infinitive. The distinction between gerunds and infinitives is a highly peculiar feature of English. In many languages, there is simply no way to distinguish between 'being dead' and 'to be dead': both expressions have the same translation in Latin (*mortuus esse*), Greek (ἀπολωλέναι), and most modern Romance languages. Furthermore, I do not anticipate that anyone will seriously suggest that dummy pronoun 'it' is doing any important work. (2) entails (3); the argument presented above is valid.

Let us consider an alternative reply to the puzzle. This reply draws inspiration from the fact that the predicate 'ϕ-ing is like ψ-ing' is highly context-sensitive. In every context, the relevant predicate expresses some sort of similarity relation. But it expresses different dimensions of similarity relative to different contexts of utterance. Thus, 'driving a real car is like driving a car in a video game' is true in a context where *rough similarity in respect of driving technique* is the salient dimension of similarity expressed by the predicate 'ϕ-ing is like ψ-ing'. But it is false in a context where *close similarity in respect of environmental friendliness* is the salient dimension of similarity expressed by the predicate 'ϕ-ing is like ψ-ing'. It makes no sense to ask whether the sentence 'driving a real car is like driving a car in a video game' is true *simpliciter*; we must ask whether it is true relative to a particular context of utterance.

With these points in mind, consider the following sentences:

1. Being super-sedated is like being in a dreamless sleep.
2. There is something that is like being super-sedated.
3. There is something that it's like to be super-sedated.

According to the present reply, (1)-(3) are true in the context of utterance ordinarily invoked by the conversation between Hal and his doctor (above). Relative to these ordinary contexts, schema 1 is false.

Nevertheless, the present reply insists that there is a *special philosophically important context of utterance* in which the predicate 'ϕ-ing is like ψ-ing' expresses a *special philosophically important similarity relation* (call it '***LIKENESS***'). Relative to this special philosophical context *C*, sentences (1)-(3) are all false. If Hal asks his doctor 'what is it like to be super-sedated?', it is ordinarily correct for the doctor to respond: 'it's like being in a dreamless sleep.' But in the special philosophical context *C*, such an answer is incorrect: being super-sedated is not ***LIKE*** being in a dreamless sleep. More generally, there is nothing it's ***LIKE*** to be super-sedated. And, most importantly: even though schema 1 is false relative to ordinary contexts of utterance, it is true relative to *C*. So the response suggests.

Unfortunately, this response scores only a hollow victory for the defender of schema 1. The response *concedes* that schema 1 is *false* on its ordinary interpretation.

Schema 1 is true, the response suggests, only if the predicate 'ϕ-ing is like ψ-ing' is interpreted in a nonstandard way, so that it becomes imbued with a meaning which it wouldn't ordinarily have. But this renders schema 1 *pointless*: its sole purpose is to capture a connection between consciousness and a natural language locution ('what it's like'), where the latter is understood in a commonsense way.

And what, exactly, is the special similarity relation of ***LIKENESS*** that the predicate 'ϕ-ing is like ψ-ing' needs to express in order to ensure the truth of schema 1? It seems extremely doubtful that an informative answer to this question is forthcoming. Surely it is better to ditch the charade that *WIL*-talk has any useful role to play in analyzing consciousness. We should take 'consciousness' as a primitive theoretical term, and decline to analyze it by appeal to any natural language locution. Or we should analyze the relevant term in some other manner.

Before we conclude, a confession is in order. Thus far, I have pretended that every relevant interpretation of the word 'like' concerns *similarity*, in some manner or other. This pretense is not entirely accurate. Questions involving the word 'like' can serve at least two functions. Suppose that Hal has never sprinted before, though he is an avid jogger, especially in high altitudes. Avril, by contrast, is a sprinter. Hal asks Avril: 'what is sprinting like?' Consider the following answers:

*A*1. It's like high-altitude jogging.
*A*2. It's exhausting but enjoyable.

Both answers are potentially acceptable, depending on the context. *A*1 answers the relevant question by identifying some activity to which sprinting is *relevantly similar* (namely, high-altitude jogging). By contrast, *A*2 does not say that sprinting is similar to anything in particular. Instead, *A*2 identifies an *interesting fact* about sprinting, a fact which has nothing to do with similarity. The differences between *A*1 and *A*2 correspond to two different purposes of 'like'-questions. One might ask 'what is sprinting like?' with a view to ascertaining some type of activity to which sprinting is similar. Or one might ask 'what is sprinting like?' with a view to learning facts about sprinting, which may or may not pertain to its similarity structure. It is natural to hypothesize that the relevant question is ambiguous; we write 'what is sprinting

like?$_1$' and 'what is sprinting like?$_2$' to express its respective disambiguations.[7, 8] When interested in similarity, one asks: what is sprinting like?$_1$ When interested in facts that needn't concern similarity, one asks: what is sprinting like?$_2$ On occasion, it may be indeterminate which question one has asked; on other occasions, one might ask both questions at once.

With these distinctions in mind, consider Hal's question: 'what is it like to be super-sedated?' Here are two perfectly good ways for Hal's doctor to answer this question:

*A*3 It's like being in a dreamless sleep.

*A*4 It's dangerous. Upon super-sedation, you will immediately become paralyzed. Your brain stem will twitch incessantly, though you won't notice.

*A*3 provides information about some state to which super-sedation is similar. *A*4, by contrast, does not concern similarity. Rather, it identifies interesting facts about super-sedation. *A*3 is a correct answer to the question 'what is it like to be super-sedated?$_1$' (or 'what is being super-sedated like?$_1$'). *A*4 is a correct answer to the question 'what is it like to be super-sedated?$_2$' (or 'what is being super-sedated like?$_2$'). Notice that there is a correct, non-negative answer to both questions. It follows that *both* of the following accounts are false:

If the question 'what is it like to ϕ?$_1$' has a positive answer, then ϕ-ing requires consciousness.

If the question 'what is it like to ϕ?$_2$' has a positive answer, then ϕ-ing requires consciousness.

---

[7] I have claimed that the *question* 'what is sprinting like?' is ambiguous. Does it follow that the *word* 'like' is ambiguous? Plausibly, yes—though I shall not pursue the matter further.

[8] On an alternative view, the relevant question is not *semantically* ambiguous, but can be employed to *pragmatically* implicate distinct questions. For present purposes, we needn't worry about the differences between these views.

After all, both 'what is it like to be super-sedated?$_1$' and 'what is it like to be super-sedated?$_2$' have positive answers, yet it is impossible for a super-sedated subject to be conscious (by the definition of 'super-sedated'). The moral is clear: the distinction between different interpretations of questions involving the word 'like' is powerless to save *WIL*-talk.

§3 A Better Way

The prospects for *WIL*-talk are dim. We have failed to uncover any plausible solution to the 'first puzzle' discussed in §**1**. (The second, third and fourth puzzles discussed in §**1** are structurally identical to the first puzzle; if there is no solution to the first puzzle, there is no solution to the others.) Fortunately, we can study consciousness without relying on *WIL*-talk. The remainder of this dissertation shows how.

There are two basic notions that any theorist of consciousness requires: the notion of phenomenal consciousness, and the notion of a phenomenal character. I propose to simply take the former as a primitive notion. In other words, I decline to analyze the notion of phenomenal consciousness by appeal to any natural language locution. (This does not rule out the possibility that further theorizing will provide an analysis of phenomenal consciousness. We can *take* a notion as primitive without assuming it to be in-principle unanalyzable.) There are many ways to get a grip on the notion of phenomenal consciousness that do not involve straight-up *definition*. We can talk about zombies (Chalmers 1996). We can talk about 'access consciousness' (Block 1995), and how it differs from phenomenal consciousness. We need not, however, use *WIL*-talk in any serious way. Phenomenal consciousness is not the first philosophical concept that eludes analysis in natural language; nor will it be the last.[9]

Continuing this theme, I propose the following definition of 'phenomenal character', which avoids any distracting usage of *WIL*-talk:

- *P* is a *phenomenal character* $=_{df.}$ *P* is a superdeterminate of consciousness.

[9] For the remainder of this dissertation, I will often shorten 'phenomenal consciousness' to simply 'consciousness'.

To explore the import of this definition, we must undertake a brief diversion into the determinable-determinate distinction.

D is a *determinable* property iff there are different ways of exemplifying *D*. If being *F* is a way of being *G*, then being *F* is a *determinate* of being *G*.[10] Various axioms uncontroversially govern the relationship between a determinable and its determinates. If *D* is a determinable property, and d_i are its determinates (where $0 \leq i \leq \alpha$, for some ordinal α), then:

- If an object exemplifies any d_i, then it exemplifies *D*.
- If an object exemplifies *D*, then it exemplifies some d_i.
- If d_i is determinate of d_j and d_j is a determinate of d_k, then d_i is a determinate of d_k. (transitivity)
- If d_i is a determinate of d_j, then d_j is not a determinate of d_i. (antisymmetry)

One paradigm determinable property is colour; other paradigm determinates include mass and shape. Increasingly-specific determinates of colour include red, scarlet, dark scarlet, and so forth. The items on this list contrast with properties such as *being red in Japan*, which is not ordinarily regarded as a 'way of being coloured' (in the relevant sense). Why not?

For one thing, being red in Japan is a more gerrymandered, less *natural* property than colour.[11] Arguably, no property is less natural than any determinable it falls under. Furthermore, it is often suggested that any two determinates of a given determinable must differ from one another in certain nontrivial respects. Thus, in his pioneering study of the determinate-determinable distinction, W.E. Johnson held that any two determinates of colour must differ from one another either in respect of hue, saturation, and/or brightness (1921: 183). Evidently, being red in Japan does not differ from any determinate of colour in these respects; so it does not count as a true determinate of colour. (See Funkhouser 2006.)

We say that *D* is a *superdeterminate* or maximally specific colour iff *D* is a colour and no colour is a determinate of *D*. Descriptions such as 'the precise shade of

[10] For a discussion of the relationship between determinable properties and their determinates with which I have sympathy, see Funkhouser (2006).
[11] In the Lewisian sense of 'natural' (1983). See CHAPTER FIVE for a discussion of naturalness.

red exemplified by this Coca-Cola can' appear to designate superdeterminate colours. It is plausible that colour space terminates in superdeterminate colours: if anything is coloured, then it exemplifies some superdeterminate colour. More generally, we say that D^* is a superdeterminate of D iff D^* is a determinate of D and no determinate of D is also a determinate of D^*. A determinable property D is said to be 'nongunky' iff every object that instantiates D also instantiates a superdeterminate of D. Colour, then, is arguably nongunky.

Let us return to (phenomenal) consciousness. It is plausible that consciousness, like colour and shape, is a determinable property. On this view, there are different ways of being conscious. Thus, there is a relatively unspecific determinate of consciousness that every subject experiencing pain exemplifies, a more specific determinate that every subject experiencing searing pain exemplifies, and so forth. Intuitively, every conscious subject is conscious in some *maximally specific way*, just as every coloured object has some maximally specific shade of colour. In the terminology introduced above, every conscious subject instantiates some superdeterminate of consciousness—a property that stands to consciousness as the precise shade of red exemplified by a particular Coca-Cola can stands to colour. I do not pretend to possess a non-question-begging argument for the view that consciousness is 'nongunky', but it is very difficult to envisage any reason for suspecting this assumption to be false. It is natural, then, to define *phenomenal characters* as the superdeterminates of consciousness.

Perhaps the most noteworthy feature of this definition is its complete avoidance of *WIL*-talk. It is also noteworthy that phenomenal characters, so understood, are properties of *conscious subjects*. (If x instantiates a determinable D, and D is nongunky, then x also instantiates some superdeterminate of D. Therefore, since consciousness is instantiated by subjects, phenomenal characters are too.) This contrasts with a common tendency in the philosophy of mind: to regard phenomenal characters as properties of *experiences*. Philosophers who quantify over experiences typically assume, implicitly or explicitly, that:

- Experiences are unrepeatable tokens: I can't enjoy the same experience twice, and you can't enjoy any of my experiences.
- Experiences are plenitudinous: my total current *perceptual* experience (e_1), my total current *visual* experience (e_2), and my current visual experience *of the computer* (e_3) all exist.
- Experiences are fine-grained: $e_1 \neq e_2 \neq e_3$.
- Experiences are mereologically related: e_3 is a proper part of e_2, and e_2 is a proper part of e_1.[12]

It is very far from obvious that there are any such things as experiences, so understood.[13] Philosophers who exclude property instances or token events from their ontology will almost certainly deny that experiences exist. And many friends of property instances or token events will doubt whether experiences can be as plenitudinous or as fine-grained as the assumptions above require. Even if experiences exist, it far from obvious that they are properly regarded as *bearers* of phenomenal characters: perhaps experiences are *instances* of phenomenal characters. In general, an instance of a property is not also a bearer of the relevant property (instances of red are not red). As such, there is little reason to suspect that experiences are bearers of phenomenal characters.

For the sake of ontological neutrality, no aspect of this thesis employs experience-talk in any substantive capacity—though it could easily be re-written to allow such talk a more central role.

§4 The Problem of Exclusion

In spite of its merits, the definition of phenomenal character provided in §**3** faces a challenge, which I dub 'the problem of exclusion'. This section explores the problem, and argues that it can be overcome.

Here is a striking fact: nothing can be both red all over and green all over. Here is another striking fact: nothing can be both square and round. Here is another

[12] See Bayne (2010) for discussion of these assumptions, and Tye (2003) for the dissenting view that experiences are sparse.
[13] Byrne (2009), for example, advocates 'experience-eliminativism'.

striking fact: nothing can both have a mass of 5g and a mass of 10g. In general, the following principle is plausible:

EXCLUSION

Same-level determinates exclude one another.

A and *B* are same-level determinates of *C* iff:

- *A* is a determinate of C;
- *B* is a determinate of *C*;
- *A* is not a determinate of *B*; and
- *B* is not a determinate of *A*.

Thus, the exclusion principle states that no object instantiates two same-level determinates of a single determinable. This principle predicts the striking observations about colour, shape and mass described above. It is widely accepted in the determinate-determinable literature: see Armstrong (1980: 112), Crane (2007: §2), Ehring (2011: 147; 1996: 470), Fine (2011), Haug (2010: 318), Johnson (1921: 181), Searle (1959), and Wilson (2009: 151).

Nevertheless, the exclusion principle conflicts with the definition of phenomenal character provided in §**3**. To see why, imagine three perceivers—Simple_1, Simple_2, and Complex. Simple_1 has a reliable visual system, but lacks any other sense modality. Simple_2 has a reliable auditory system, but lacks any other sense modality. In contrast, Complex has a reliable auditory system and a reliable visual system which resemble those of Simple_1 and Simple_2 in relevant respects. Imagine that Simple_1 stands in front of a radio emitting a constant noise. By virtue of visually perceiving the radio, Simple_1 instantiates a phenomenal character P_1. Shortly thereafter, Simple_2 stands in front of the same radio from the same perspective. By virtue of auditorily perceiving the radio, Simple_2 instantiates a phenomenal character P_2. Finally, Complex stands in front of the same radio from the same perspective. Complex is fortunate enough to see and hear the radio. Intuitively, Complex instantiates *both* P_1 and P_2. More generally, it is plausible that subjects with more than one sense modality instantiate more than one phenomenal character, regardless of how the notion of 'phenomenal character' is ultimately defined.

Unfortunately, this intuition conflicts with the view that phenomenal characters are superdeterminates of consciousness. The exclusion principle entails that nothing can have two superdeterminates of the same determinable, for all such superdeterminates are at the 'same level'. If phenomenal characters are superdeterminates, it follows that nothing can have two phenomenal characters—contrary to the plausible claims advanced above.

Of course, this problem only arises if phenomenal characters are defined as superdeterminates of phenomenal consciousness. Should we abandon the proposed definition?

I shall outline three responses to the problem outlined above. Only the third response will prove compelling.

First response: brains and consciousness

Here is an initial suggestion. One part of Complex's brain instantiates P_1. Another part of Complex's brain instantiates P_2. But no single part of Complex's brain instantiates both P_1 and P_2. On this view, no single thing instantiates two phenomenal characters: Complex himself instantiates neither P_1 nor P_2. No one thinks that multi-coloured objects violate the exclusion principle; rather, different parts of such objects instantiate different colours. On this view, Complex is rather like a multi-coloured object.[14]

Thus far, I have regarded *subjects* as the bearers of phenomenal consciousness (and hence phenomenal characters). The reply considered above conflicts with this construal: it requires us to regard phenomenal consciousness as a property instantiated by *parts of brains*, and not by subjects such as Complex. This view is reminiscent of a general trend in cognitive science, whereby brains and parts of brains are described as having beliefs and other mental states. What should we make of such talk?[15]

We might react in a hostile manner, and hold that all ascriptions of beliefs to brains (and parts thereof) are simply false, understood literally. Alternatively, a conciliatory reaction is available. We might agree that brains (and parts thereof) can

[14] A similar response is available to those who regard *experiences* as the bearers of phenomenal properties (cf. §3). One experience instantiates P_1; a different experience instantiates P_2.

[15] One might regard the distinction between brains and subjects as null if one identifies subjects with their brains. But there are decisive objections against this position (Olson 2007).

have beliefs, but deny that brains are *fundamental* bearers of beliefs. That is: if a brain believes that p, it does so only by virtue of standing in a salient relation to a person or other organism who believes that p. A third, revisionist reaction is also possible: one might hold that brains are among the fundamental bearers of beliefs.

It is easy to have sympathy with the hostile and conciliatory reactions, and very difficult to have sympathy with the revisionist reaction. The reasons for this are complex, and there is no space to consider the matter carefully. Nevertheless, I suspect that many readers will share these sentiments.

So far, we have merely considered belief. But it seems to me that matters are much the same in the case of phenomenal consciousness. Perhaps it is literally false to describe a brain (or a part thereof) as phenomenally conscious. Or perhaps brains (or parts thereof) are phenomenally conscious only by virtue of their relation to a phenomenally conscious person. Both of these claims conflict with the response considered above, which holds that Complex's brain is phenomenally conscious even though Complex himself is not. (After all, the response claims that Complex himself does not instantiate any phenomenal characters, and x is phenomenally conscious only if x has some phenomenal character.) The response is objectionable for precisely this reason.

Second response: phenomenal holism

Let us consider a second response to the problem of exclusion. This response also denies that Complex instantiates both P_1 and P_2. Instead, Complex is said to instantiate a complex phenomenal character P_3, which somehow *combines* the simpler phenomenal characters P_1 and P_2. Although P_3 is (in some sense) a 'combination' of P_1 and P_2, the reply under consideration denies that subjects who instantiate P_3 must also instantiate the phenomenal characters of which it is combined (viz. P_1 and P_2). It is natural to regard this view as a form of *phenomenal holism*. The opposing view—that Complex instantiates both P_1 and P_2—may be regarded as a form of *phenomenal atomism*.

Phenomenal holism is an unattractive position. I have no objection to the view that Complex instantiates a complex phenomenal character P_3 which somehow combines P_1 and P_2. But it is strange, I submit, for phenomenal holists to insist that a subject can instantiate P_3 without also instantiating the phenomenal characters of

which it is combined (P_1 and P_2). After all, what does it mean to say that P_3 'combines' P_1 and P_2? The most natural reading of this claim is that P_3 is a *conjunctive* property whose conjuncts are P_1 and P_2. But no object can instantiate a conjunctive property without instantiating its conjuncts: indeed, to instantiate a conjunctive property *just is* to instantiate its conjuncts.

Phenomenal holists may deny that the relevant notion of 'combination' can be analyzed in terms of the notion of a conjunctive property. But this renders the relevant notion mysterious. Alternatively, phenomenal holists may withdraw the claim that P_3 is a 'combination' of P_1 and P_2. But this renders phenomenal holism unmotivated. Phenomenal holism initially purported to *explain away* the intuition that Complex instantiates both P_1 and P_2, by appealing to the idea that Complex instantiates a distinct phenomenal property P_3 which *combines* P_1 and P_2. In the absence of this explanation, phenomenal holism is nothing more than an unmotivated denial of a plausible intuition.

Third response: denying exclusion

A third (and final) response is to deny the principle of exclusion. On this view, Complex instantiates at least two superdeterminate phenomenal characters, P_1 and P_2. The exclusion principle—which states that all same-level determinates of the same determinable exclude one another—is therefore false.

This is my preferred response to the problem of exclusion. The principle of exclusion has a venerable history: Armstrong (1980: 112), Crane (2007: §2), Ehring (2011: 147; 1996: 470), Fine (2011), Haug (2010: 318), Johnson (1921: 181), Searle (1959), and Wilson (2009: 151) all accept it. To the best of my knowledge, no one has ever denied the principle of exclusion in print. Nevertheless, the principle is threatened by a number of persuasive counter-examples.

Determinable	**Determinate (1)**	**Determinate (2)**
Having a child	Having a son	Having a daughter
Being a scientist	Being a physicist	Being a biologist
Being an artist	Being a musician	Being a painter
Being a pet-owner	Being a dog-owner	Being a cat-owner

The table above gives four examples of determinable properties with same-level determinates that do not exclude one another. (If you don't like all the examples, choose your favourite.)

Proponents of the exclusion principle will presumably reply that none of the examples above are genuine determinables. In the absence of further explanation, this response is objectionably *ad hoc*. There is a clear sense in which there are different ways of having children, different ways of being a scientist, different ways of being an artist, and different ways of being a pet-owner. The exclusion principle is the result of excessive emphasis on colour, and too little focus on other types of determinable properties. To put a recursive spin on the same point: perhaps there is a *determinate* of the *determinate-of* relation (a way of being a way of being) which satisfies the exclusivity principle. But the principle does not universally hold.

…

In summary, I have advocated the following claims:

(*i*) Theorists investigating consciousness should avoid *WIL*-talk.
(*ii*) Instead, we should take 'phenomenal consciousness' as primitive, and define phenomenal characters as superdeterminates of phenomenal consciousness.
(*iii*) This account faces a problem: the problem of exclusion. But the problem of exclusion can be overcome.

§5 Changing Gear: Representationalism

My central projects in the present chapter are now complete. Before moving on, it will prove fruitful to develop an official formulation of one leading account of perceptual phenomenology: *representationalism*. In subsequent chapters, I shall argue that representationalism conflicts with our best theory of vagueness.

Prior to devising an official formulation of representationalism, some background is required. Let us begin with the following Harman-esque insight: *perceiving the world involves perceptually predicating properties of objects in one's*

environment.[16] For example: Avril perceptually predicates redness of a tomato (or perhaps some specific shade of redness); Hal perceptually predicates circularity (or some nearby shape) of a donut. In addition to relatively simple properties like redness and circularity, perceivers often perceptually predicate very complex properties. Imagine, for example, that Hal is viewing a red square and a blue circle; Avril is viewing a blue square and a red circle. Both Avril and Hal perceptually predicate redness, blueness, squareness, and circularity. However, Hal perceptually predicates various complex properties, such as:

1. $\lambda x \lambda y(\text{red}(x) \ \& \ \text{square}(x) \ \& \ \text{blue}(y) \ \& \ \text{circular}(y))$

(1) employs the helpful and familiar tool of λ-abstraction. Informally, the expression '$\lambda x \lambda y(\text{red}(x) \ \& \ \text{square}(x) \ \& \ \text{blue}(y) \ \& \ \text{circular}(y))$' denotes a complex dyadic property that an object *x* and an object and *y* instantiate iff *x* is red and square, and *y* is blue and circular. This is admittedly a rather gerrymandered property—but believers in plenitudinous properties should have no objection to its existence. (On plenitudinous properties, see INTRODUCTION note 1.)

By contrast, Avril perceptually predicates a different complex property:

2. $\lambda x \lambda y(\text{red}(x) \ \& \ \text{circular}(x) \ \& \ \text{blue}(y) \ \& \ \text{square}(y))$

Informally, the expression '$\lambda x \lambda y(\text{red}(x) \ \& \ \text{circular}(x) \ \& \ \text{blue}(y) \ \& \ \text{square}(y))$' denotes a complex dyadic property that an object *x* and an object and *y* instantiate iff *x* is red and circular, and *y* is blue and square.

As these remarks indicate, I regard the two-place predicate: '*x* perceptually predicates *y*' (where *x* is a perceiver and *y* is a property) as the distinctive theoretical ideology of representationalism. I shall also employ the three-place predicate: '*x* perceptually predicates *y* of $<z_1 \ldots z_n>$', where $<z_1 \ldots z_n>$ is a sequence of objects which *x* perceives.[17] The locutions are, of course, connected: if there exist objects $z_1 \ldots z_n$ such

[16] See Harman (1990). The notion of perceptual predication is discussed by many writers. Among the best recent treatments are Siegel (2010) and Pautz (2009).
[17] For present purposes, a sequence of length 1 = $<z>$ is simply identified with *z*; a sequence of length $n>1 = <z_1 \ldots z_n>$ is identified with the corresponding Kuratowski *n*-tuple. See any introductory set theory textbook for details.

that x perceptually predicates y of $<z_1 \ldots z_n>$, then x perceptually predicates y. However, the converse arguably fails: in cases of 'perfect hallucination', perceivers perceptually predicate properties without perceptually predicating them *of* anything. For convenience, I shall often shorten 'perceptual predication' to the less-cumbersome 'predication'.

Representationalists almost invariably decline to analyze predication by appeal to any natural language locution, taking particular care to distinguish ordinary expressions such as 'x looks F' and from their notion of predication. However, representationalists typically hold that there is a more-or-less pre-theoretical notion of *perceptual accuracy* which is directly connected to predication:

- If A is perceiving $B_1 \ldots B_n$, then A perceives $B_1 \ldots B_n$ accurately iff the sequence $<B_1 \ldots B_n>$ is in the extension of every property which A predicates of $<B_1 \ldots B_n>$.

The notion of perceptual predication gives rise to many questions. It is uncontroversial that colour properties and spatial properties are regularly predicated. Is it possible to predicate natural kind properties, such as *being a fish*? What about artifact properties such as *being a table*? There is a large and illuminating literature discussing such questions. For present purposes, however, we can these issues aside.

According to representationalism, there is an intimate connection between perceptual predication and perceptual phenomenology. More carefully: facts about perceptual predication and facts about perceptual phenomenology stand in a relation of co-determination. More carefully still: say that x is a *perceptual phenomenal character* iff x is a phenomenal character, and it is impossible to instantiate x at a time t without perceiving at t. Many phenomenal characters are perceptual. But some—for example, the phenomenal characters involved in imagination—are plausibly nonperceptual. According to representationalism,

REPRESENTATIONALISM

For every perceptual phenomenal character x, there is some property y such that: a subject instantiates x iff she perceptually predicates y.[18]

$\forall x$(perceptual-phenomenal(x) $\rightarrow \exists y \forall z$($z$ instantiates $x \leftrightarrow z$ perceptually predicates y)

To put the same point in different language, representationalists claim that there is an injective function f from the class of perceptual phenomenal characters into the class of properties, such that a subject exemplifies a perceptual phenomenal character x iff she predicates $f(x)$. Such a function is depicted in Figure 1, below.

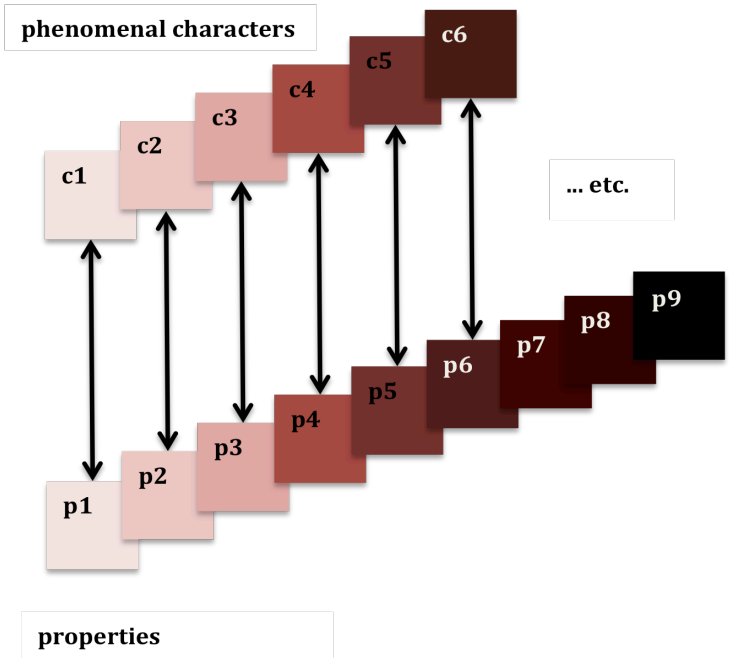

Figure 1

In general, the predicable property associated with a given perceptual phenomenal character will be extremely complex, similar to the properties expressed using λ-abstraction above.

Representationalism is the most popular account of perceptual phenomenology in contemporary philosophy.[19] Some representationalists (e.g. Tye

[18] This is a strict biconditional: *necessarily*, a subject instantiates x iff she predicates y.
[19] Proponents of representationalism (or something near enough) include Byrne (2001), Dretske (1995), Johnston (2004), Harman (1990), Hill (2009), Pautz (2007) and Tye (1995; 2000).

1995 and Dretske 1995) are motivated by reductive ambitions, driven by the hope that consciousness can be naturalized through a causal-covariational account of perceptual predication. Others (e.g. Pautz 2006) have no reductive ambitions whatever, and simply regard representationalism as an interesting thesis about the relationship between perceptual phenomenal character and predication. I suspect that many proponents of 'naïve realism' also subscribe to representationalism (when formulated as above), even though naïve realism is often construed as an *alternative* to representationalism. For more on this theme, see Siegel (2010).

It goes without saying, of course, that there are many alternative formulations of representationalism in recent literature. I do not intend to undertake the tedious task of distinguishing every such formulation from the one provided above; there is already a cottage industry devoted to addressing such matters. I will only distinguish my full-strength version of representationalism from three different formulations.

WEAK REPRESENTATIONALISM

For any perceptual phenomenal character *x*, some property *y* is such that:
a subject instantiates *x* if she predicates *y*.

Comment: weak representationalists are only committed to one direction of the biconditional implied by full-strength representationalism. (The 'if' direction, not the 'only if' direction.) The arguments developed in this dissertation *do not* target weak representationalism; more on this in CHAPTER TWO §**4**.

VISUAL REPRESENTATIONALISM

For any visual phenomenal character *x*, some property *y* is such that:
a subject instantiates *x* iff she visually predicates *y*.

Comment: visual representationalism exchanges the generic notion of perceptual predication for a more specific notion of visual predication, and the generic notion of a perceptual phenomenal character for the more specific notion of a visual phenomenal character. Other modality-specific versions of representationalism are easily envisaged. Every argument provided below applies *mutatis mutandis* to visual representationalism.

PROPOSITIONAL REPRESENTATIONALISM

For any perceptual phenomenal character x, some proposition y is such that:
a subject instantiates x iff she perceptually represents y.

Comment: propositional representationalism exchanges the predicate 'x perceptually predicates y' (which takes perceivers and properties as relata) for the alternative predicate 'x perceptually represents y' (which takes perceivers and propositions as relata). Presumably, a subject perceptually represents the following *proposition*:

$\exists x \exists y(\text{red}(x) \ \& \ \text{square}(x) \ \& \ \text{blue}(y) \ \& \ \text{circular}(y))$

iff she perceptually predicates the following complex *property*:

$\lambda x \lambda y(\text{red}(x) \ \& \ \text{square}(x) \ \& \ \text{blue}(y) \ \& \ \text{circular}(y))$

As such, it is plausible that propositional representationalism is equivalent to my formulation of representationalism. For more on this theme, see Pautz (2007).

Finally: in what follows, I will reserve the term 'phenomenal character' for *perceptual* phenomenal characters, unless otherwise stated.

Representationalism and the Problem of Vagueness

CHAPTER TWO

This chapter develops a new problem for the representationalist theory of perceptual phenomenology.[20] The problem—in a nutshell—is this. Given plausible assumptions, representationalism is incompatible with the following claims:

A. Perceptual predication is vague.
B. Phenomenal characters are precise.

As we shall see, (A) and (B) are directly implied by our best theory of vagueness: *precise supervaluationism*. In effect, then, this chapter argues that recent trends in the philosophy of perception and the philosophy of vagueness conflict with one another. This is the 'problem of vagueness'.

As it stands, (A) and (B) are mere slogans; their contents are unpacked in a rigorous manner later on. Nevertheless, even in their inchoate present condition, the potential for incompatibility with representationalism should be clear: representationalism claims that there is an intimate connection between phenomenal character and perceptual predication; (A) and (B) suggest otherwise.

The big question, of course, is whether any satisfactory solution to the problem of vagueness is available to card-carrying representationalists. I do not pretend to offer an authoritative answer to this question. Indeed, I am inclined to believe that the problem of vagueness runs fairly deep; 'cheap solutions' are not forthcoming. There seem to be only two serious responses available to the representationalist, and neither is particularly attractive. Whether either response is ultimately successful turns on subtle issues pertaining to the nature of vagueness which I cannot hope to fully explore here. My central goal is to encourage discussion of these under-explored issues, not to solve them.

[20] I have explored the problem elsewhere in [reference omitted for anonymous review].

We shall proceed as follows. §**1** introduces several basic tools for theorizing about vagueness. **§2** discusses precise supervaluationism. **§3** argues that precise supervaluationism implies (A) and (B). **§4** proves (in a classical setting) that representationalism is inconsistent with (A) and (B), provided that a *prima facie* plausible assumption is granted. This concludes the present chapter. Discussion of putative resolutions to the problem of vagueness will occur in the subsequent two chapters.

§1 Vagueness: The Basics

The concept of vagueness is best introduced by appeal to the concept of a *borderline case*. In this connection, Cian Dorr writes:

> The concept [of a borderline case] has its most basic application when we are faced with a question of the form 'Is *x F*?', but are unwilling to answer 'yes' or 'no' for a certain kind of distinctive reason. Wanting to be co-operative, we need to say something; by saying 'it's a borderline case', we excuse our failure to give a straightforward answer while conveying some information likely to be of interest to the questioner.[21]

Examples are plenitudinous. Is Hal bald? It's a borderline case. Is Avril old? It's a borderline case. Is a particular patch of colour an instance of red or of orange? It's a borderline case. Is Orin tall? It's a borderline case. Such 'borderline case'-talk is commonly regimented using the one-place sentential operator 'it is vague whether…' Thus, we say:

1. It is vague whether Hal is bald.
2. It is vague whether Avril is old.
3. It is vague whether Orin is tall.

[21] Dorr 2010: 550.

In formal languages, the operator 'it is vague whether…' is customarily replaced with the less-cumbersome operator ∇.[22] Thus, 'it is vague whether Hal is bald' is rendered: ∇Bh.[23]

The concept of vagueness is connected to the concept of *determinacy*. It is determinate that an ordinary man with 3 hairs on his head is bald. It is determinate that a three-foot grown woman is not tall. It is determinate that 5,000,000,000 grains of sand placed closely together compose a heap. It is determinate that scarlet is a shade of red. When theorizing in a formal language, the operator 'it is determinate that…' is customarily replaced with the less-cumbersome operator Δ. Thus, 'it is determinate that Hal is bald' is rendered: ΔBh.[24] The connection between vagueness and determinacy is captured by the following schemata:

$\nabla\phi \leftrightarrow (\sim\Delta\phi \ \& \ \sim\Delta\sim\phi)$

It is vague whether ϕ iff (it is neither determinate that ϕ, nor determinate that not-ϕ).

$\Delta\phi \leftrightarrow (\phi \ \& \ \sim\nabla\phi)$

It is determinate that ϕ iff (ϕ and it is not vague whether ϕ).

If you understood vagueness, you now understand determinacy.

---

[22] See Williamson (1999), among many others.

[23] Some philosophers (e.g. Eklund 2011) distinguish *indeterminacy* from *vagueness*. These philosophers claim that, on their intended interpretation of 'indeterminate', the following claims are true:

1. It is indeterminate whether the Liar sentence is true, but it is not vague whether the Lair sentence is true.
2. It is indeterminate whether a certain coin will land heads when flipped tomorrow, but it is not vague whether the relevant coin will land heads.

This is not *my* usage of 'indeterminate': I will treat 'indeterminate' as synonymous with 'vague' and 'borderline'. (That is not to say, of course, that Eklund's usage is incoherent; perhaps there is a notion of indeterminacy that has nothing to do with the notion of a borderline case.)

[24] It is worth distinguishing the locution 'it is determinate that ϕ' from the locution 'it is determinate *whether* ϕ'. The locutions are interdefinable: it is determinate whether $\phi \leftrightarrow (\Delta\phi \vee \Delta\sim\phi)$. Likewise, $\Delta\phi \leftrightarrow$ (it is determinate whether ϕ) & ϕ.

Thus far, we have regimented talk about vagueness and determinacy by using *sentential operators* such as ∇ (= 'it is vague whether…') and Δ (= 'it is determinate that…'). However, for stylistic reasons, I will occasionally theorize about vagueness and determinacy informally without using sentential operators. For example, it is often convenient to place 'determinately' between the copula and an adjective, as in:

4. Hal is determinately bald.

Fortunately, it is straightforward to translate such claims into the formal language of sentential operators. I hereby stipulate that (4) is rendered thus: ΔBh. All other nonstandard vagueness-theoretic locutions that play any important role in this chapter will always be accompanied by a translation into a formal language using Δ and ∇.[25]

Before we proceed further, three notes on logical matters. First, every argument in this dissertation has used, and will continue to use, classical propositional and first-order logic. Some philosophers deny that classical logic has any useful application to vague languages. There is little I can do to assuage these philosophers' concerns. The relationship between classical logic and vagueness is a truly enormous topic, far beyond the purview of the present dissertation. Nevertheless, it is now widely recognized that classical logic provides an extremely valuable framework for reasoning about vagueness, one that cannot easily be replaced.[26]

Second, I will assume that Δ obeys a *weak modal logic* (Williamson 1999). Such a logic comprises two axiom-schemata, (K_Δ) and (T_Δ), and one inference-rule, (N_Δ):

(K_Δ) $\Delta(\phi \rightarrow \psi) \rightarrow (\Delta\phi \rightarrow \Delta\psi)$

If a conditional is determinate, and its antecedent is also determinate, then so is its consequent.

[25] A small number of philosophers argue that it is somehow *bad* to theorize about vagueness using the sentential operators Δ and ∇. See Dorr (2010) for an argument against this position.

[26] As Robbie Williams observes in a recent draft (ms: 1), 'classical treatments of indeterminacy are on the march. A growing number of authors argue that we can have an adequate theory of indeterminacy or vagueness that demands no revision of the classicism-presupposing theories used throughout the sciences.'

(T_Δ) $\Delta\phi \rightarrow \phi$

If it is determinate that ϕ, then ϕ.

(N_Δ) If ϕ is a logical consequence of (K_Δ) and (T_Δ), then $\Delta\phi$ is true.

(K_Δ), (T_Δ), and (N_Δ) correspond to the weakest logic for the necessity operator $\Box$:

($K_\Box$) $\Box(\phi \rightarrow \psi) \rightarrow (\Box\phi \rightarrow \Box\psi)$

If a conditional is necessary, and its antecedent is also necessary, then so is its consequent.

($T_\Box$) $\Box\phi \rightarrow \phi$

If it is necessary that ϕ, then ϕ.

($N_\Box$) If ϕ is a logical consequence of ($K_\Box$) and ($T_\Box$), then $\Box\phi$ is true.

It is extremely commonplace to use weak modal logics when reasoning with both $\Box$ and Δ. There is a straightforward reason for this. Without such logics, it is impossible to prove anything interesting about either necessity (/possibility) or determinacy (/vagueness). One particularly useful theorem of the weak modal logic for Δ, to which we will regularly appeal, is this:

(V) $\Delta(\phi \leftrightarrow \psi) \rightarrow (\nabla\phi \leftrightarrow \nabla\psi)$

If a biconditional is determinate, then its antecedent is vague iff its consequent is vague.

The proof of (V) is relegated to a footnote.[27]

---

[27] *Assumption A*: $\Delta(\phi \leftrightarrow \psi)$.

1.	$\Delta\phi \leftrightarrow \Delta\psi$	(assumption A, K_Δ)
2.	$(\phi \leftrightarrow \psi)$	(T_Δ, assumption A)
3.	$(\sim\phi \leftrightarrow \sim\psi)$	(classical logic, 2)
4.	$\Delta\sim\phi \leftrightarrow \Delta\sim\psi$	(K_Δ, 3)

Assumption B (for *reductio*): $\nabla\phi$ & $\sim\nabla\psi$.

5.	$\Delta\psi \vee \Delta\sim\psi$	(assumption B, definition of Δ)
6.	$\Delta\psi \rightarrow \Delta\phi$	(modus ponens, 1)
7.	$\Delta\psi \rightarrow \bot$	(assumption B, 6)
8.	$\Delta\sim\psi \rightarrow \Delta\sim\phi$	(modus ponens, 4)
9.	$\Delta\sim\phi \rightarrow \bot$	(assumption B, 8)
10.	$(\Delta\psi \vee \Delta\sim\psi) \rightarrow \bot$	(disjunction elimination: 5, 7, 9)

[*cont'd next page*]

Finally, we shall *not* rely on any instance of the following schema:

(TRIV) $\phi \rightarrow \Delta\phi$

If ϕ, then it is determinate that ϕ.

Such neutrality has much to recommend it. Most philosophers of vagueness deny certain instances of (TRIV). By contrast, I have often witnessed philosophers of mind in conversation express sympathy with the view that every instance of (TRIV) is true. These sympathizers generally regard the notion of an 'indeterminate truth', to which violations of (TRIV) would give rise, as absurd. In any case, no argument at any stage of this chapter is reliant on any instance of (TRIV).

§2 Precise Supervaluationism

The reader is doubtless familiar with many accounts of vagueness that parade under the banner of 'supervaluationism'. It is no exaggeration to state that supervaluationism is *the* dominant semantics for Δ and ∇ in contemporary philosophy. I shall explore a particular type of supervaluationism—*precise supervaluationism*, in my jargon—which has proven particularly popular.[28] As we shall see, precise supervaluationism gives rise to serious difficulties for the representationalist theory of perception.

Let us begin with the following example. It is vague whether 20 is a small number. (If you don't like 20, choose some other number. The predicate '*x* is a small number' is extremely context-sensitive, so there is nothing wrong with assuming that

---

11. $\sim(\nabla\phi \ \& \sim\nabla\psi)$ (*reductio*: assumption B, 5, 10)

Hence,

12. $\Delta(\phi \leftrightarrow \psi) \rightarrow (\nabla\phi \rightarrow \nabla\psi)$

The other direction:

13. $\Delta(\phi \leftrightarrow \psi) \rightarrow (\nabla\psi \rightarrow \nabla\phi)$

is proven in the same way.

[28] Why 'precise supervaluationism', instead of simply 'supervaluationism'? The answer to this question will eventually become apparent. In brief: the term 'supervaluationism' is occasionally reserved for a mere formal semantics or model theory (in effect, a mere algebra) which lacks substantive implications. 'Precise supervaluationism' is *not* mere model theory.

we are in a context in which it is vague whether 20 counts as small.) To keep things simple, let us also assume that hardcore Platonism is true: there exists some object which is determinately identical to the number 20.

Here is a *bad view* about vagueness. There is a particular property *V* which the predicate '*x* is a small number' determinately expresses. It is vague whether 20 is a small number because it is vague whether 20 instantiates *V*. Picture the matter thus:

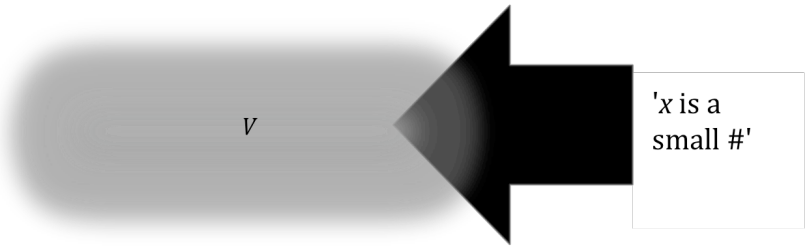

Figure 2: *V* is a 'vague property', represented by the fuzzy grey blob. The predicate '*x* is a small number' determinately expresses the relevant property, as the sharp black reference-arrow indicates.

I propose to simply assert that the view depicted by Figure 2 is a bad view. There is something very queer about vague properties; we should exclude them from our ontology. (Substantive arguments are provided later on.)

Here is a *better view*. There is no particular property which the predicate '*x* is a small number' determinately expresses, and no property *V* such that it is vague whether 20 instantiates *V*. Rather, there are many precise properties in the vicinity, and it is vague which of them is expressed by the predicate '*x* is a small number':

- the property of being less than 18;
- the property of being less than 19;
- the property of being less than 20;
- the property of being less than 21;
- etc.

We might call these properties *small-candidates*. *y* is a small-candidate iff it is vague whether '*x* is a small number' expresses *y*. Notice that every small-candidate is completely precise: every number is either determinately less than 18 or determinately

not less than 18, determinately less than 19 or determinately not less than 19, and so forth. Picture the matter thus:

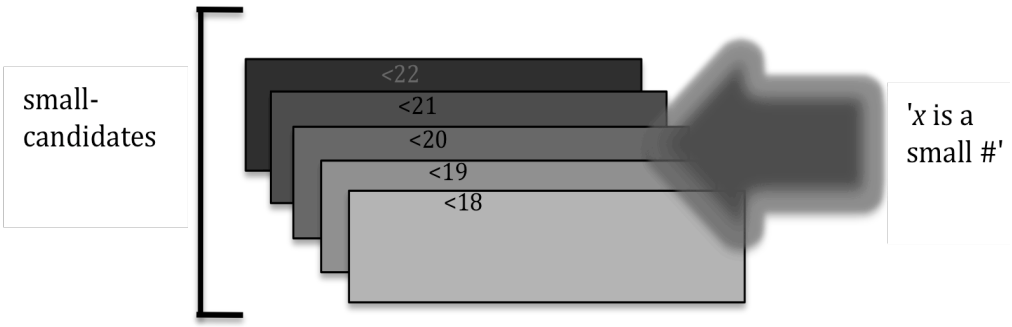

Figure 3: the rectangles represent small-candidates: the property of being less than 18, 19, and so forth. It is vague which small-candidate is expressed by the predicate '*x* is a small number', as the fuzzy reference-arrow indicates.

Of course, not every property is a small-candidate: the property of being greater than 1 is not a small-candidate; nor is the property of being less than a strongly inaccessible cardinal number.

Crucially, *it is vague whether 20 is a small number iff 20 instantiates some (but not all) small-candidates*. Thus, 20 instantiates the property of being less than 21, but does not instantiate the property of being less than 18. Since these properties are both small-candidates, it follows that it is vague whether 20 is a small number.

As noted above, every small-candidate is precise. There is no fuzzy grey blob in the ontology depicted by Figure 3; every property is a crisp rectangle. Accordingly, the following inference is invalid:

It is vague whether α is F.
Therefore, there is some property x such that it is vague whether α instantiates x.

From the fact that it is vague whether α is F, it follows merely that there are various F-candidates—some of which α determinately instantiates, and the rest of which α determinately does not instantiate. It does not follow that there is some 'vague property' V such that it is vague whether α instantiates V.

That, in a nutshell, is the basic idea behind precise supervaluationism. The same idea applies across all cases of vagueness. 'Bald-candidates' are precise

properties pertaining to the distribution of hair. It is vague whether Tom is bald because Tom instantiates some (but not all) bald-candidates. 'Red-candidates' are precise properties which correspond to precise regions of colour space. It is vague whether a given shade is red because it instantiates some (but not all) red-candidates. And so forth. Let us provide a more rigorous (and general) formulation of precise supervaluationism. We begin with the following definitions:

- For any property x and any predicate F: x is an <u>F-candidate</u> iff either F determinately expresses x, or it is vague whether F expresses x.
  Formally: $(\Delta(F \text{ expresses } x) \vee \nabla(F \text{ expresses } x))$
- When there is more than one F-candidate, we say that it is *vague which* property F expresses.

- A property x is <u>precise</u> iff necessarily, there is no object y such that it is vague whether y instantiates x. Otherwise, x is <u>vague</u>.
  Formally: $\forall x(x \text{ is a property} \rightarrow (\text{precise}(x) \leftrightarrow \Box{\sim}\exists y\nabla(y \text{ instantiates } x)))$

Precise supervaluationism comprises two claims.

PRECISE SUPERVALUATIONISM

- Every property is precise.
- For any monadic predicate F and any object x: it is vague whether F applies to x iff x instantiates some (but not all) F-candidates.[29]

The idea, then, is that every vague predicate F is associated with a variety of precise properties or 'candidates' $y_1 \ldots y_n$ such that it is vague which y_i is expressed by F. Vagueness arises when an object instantiates some (but not all) y_i. As we observed above, it is vague whether ϕ iff it is neither determinate that ϕ, nor determinate that not-ϕ. Thus, precise supervaluationism implies that:

[29] Likewise for polyadic predicates: for any objects $x_1 \ldots x_n$, it is vague whether an n-place predicate F applies to an n-tuple $<x_1 \ldots x_n>$ iff $<x_1 \ldots x_n>$ is in the extension of some (but not all) F-candidates.

- For any monadic predicate F and any object x: it is determinate that F applies to x iff x instantiates every F-candidate.

The number 0, for example, is determinately small, since it instantiates every small-candidate.

It is important to stress that the vast majority of precise supervaluationists accept the following claim:

1. Some property is expressed by 'small'.
   $\exists x$(x is expressed by 'small')

Precise supervaluationists merely hold that it is vague *which* property is expressed by the relevant word. Thus, they deny the following *de re* claim:

2. Some property is determinately expressed by 'small'.
   $\exists x \Delta$(x is expressed by 'small')

By accommodating vagueness without admitting vague properties, precise supervaluationism captures the familiar suggestion that 'the world itself' does not suffer from vagueness. Vagueness arises only when our representations fail to determinately latch onto any particular chunk of the world. Thus, Michael Dummett writes:

> The notion that things might actually be vague, as well as being vaguely described, is not properly intelligible.[30]

Likewise, Bertrand Russell writes:

> There is a certain tendency in those who have realized that words are vague to infer that things are also vague… This seems to me precisely a case of the fallacy of verbalism—the fallacy that consists in mistaking the properties of

[30] Dummett 1975: 260.

> words for the properties of things. Vagueness and precision alike are characteristics which can only belong to a representation […][31]

In the same vein, David Lewis writes:

> 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 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.[32]

Other proponents of precise supervaluationism include Kit Fine (1975), Vann McGee and Brian McLaughlin (1995), Kirk Ludwig and Greg Ray (2002), David Braun and Ted Sider (2007), Timothy Williamson (2003), and many, many others. As Williamson's inclusion on the latter list demonstrates, precise supervaluationism is compatible with (and perhaps even entailed by) certain versions of epistemicism: see Williamson (*ibid*).

Before we move on, two brief observations. First, this dissertation does not assume that ⌜ϕ⌝ is true only if it is determinate that ϕ, as some (but certainly not all) supervaluationists have claimed.[33] Second, the version of supervaluationism developed in this chapter falls far short of providing a complete supervaluationist semantic theory for a vague first-order language. In my judgment, however, there is little to be gained from expounding the minutiae of such a theory; readers are instead advised to consult the works cited above.

We shall have much more to say about precise supervaluationism later on. Nevertheless, the view should be clear enough to work with for present purposes.

---

[31] Russell 1923: 83.
[32] Lewis 1986: 212-13.
[33] This claim (sometimes expressed as the view that 'truth = supertruth') has become increasingly unpopular since McGee and McLaughlin's classic 1995 paper.

§3 Precise Supervaluationism and Perception

This section argues that precise supervaluationism implies the following claims:

A. Perceptual predication is vague.
B. Every phenomenal character is precise.

To keep things simple, let us focus on a particular perceiver—call her 'Avril'. In what follows, 'Avril' is treated as a free variable standing for an arbitrary perceiver. Reasoning with variables standing for arbitrary objects is common throughout mathematics and philosophy, invaluable for proving existential and universal truths (Fine 1985), so I hope there is no objection to this procedure. Avril is viewing a medium-sized canvas (call it 'β'), painted a uniform shade of dark red. β is located one foot in front of Avril, but lies in the periphery of her visual field. For simplicity, we shall assume that Avril does not perceive any other object. This assumption could easily be eliminated, but it enables us to avoid needless epicycles. Avril is an ordinary perceiver, viewing β under normal conditions. Presumably, therefore, it is *determinate* that she perceives β accurately.

Nevertheless, it could easily have been *vague* whether Avril perceives β accurately. To see that this is so, suppose that the superdeterminate shade of dark red that β actually has is R_{500}. Let $<w_@, w_1, \ldots, w_k>$ be a finite sequence of possible worlds, where $w_@$ is the actual world. β has very slightly different shades of colour at every world in the sequence. Thus, β has R_{500} in $w_@$. β has R_{499} in w_1. β has R_{498} in w_2. β has R_{497} in w_3. And so forth. In w_k—the final world in the sequence—β is no longer dark red, but some maximally specific shade of bright orange. However, let us stipulate that Avril perceives β in exactly the same way at every world in the sequence. That is: she perceptually attributes exactly the same properties to β, and instantiates exactly the same phenomenal characters, in every world. The canvas changes colour from world to world; Avril's perception of the canvas remains constant.

It should be clear that $<w_@, w_1, \ldots, w_k>$ forms a Sorites sequence. It is determinate that Avril perceives β accurately in $w_@$. Her perceptual system is functioning normally, and she is viewing β under ordinary conditions. Likewise, it is

determinate that Avril *does not* perceive β accurately in w_k: β is bright orange in w_k, yet Avril continues to perceptually attribute some shade of red to β. Nevertheless, there are many worlds w_i, positioned near the middle of the Sorites sequence, such that it is *vague* whether Avril perceives β accurately in w_i. I suspect that many readers will (correctly) regard this claim as a banal truth. Not convinced? Notice that if it is not vague whether Avril perceives β accurately at any world in the sequence, it immediately follows that:

For some world w_i: it is determinate that Avril perceives β accurately in w_i, and it is determinate that Avril does not perceive β accurately in w_{i+1}.

Surely there is no such 'determinate cut-off point' in the series of worlds $<w_@, w_1, \ldots, w_k>$, any more than there is a determinate cut-off point separating small numbers from large numbers. There must, therefore, be various worlds w_i such that it is vague whether Avril perceives β accurately in w_i, just as there are various numbers n such that it is vague whether n is small.[34]

β changes colour across $<w_@, w_1, \ldots, w_k>$. However, there is nothing special about colour; we can set up the Sorites series in various alternative ways and arrive at the same result. For example, let us stipulate that β changes its *location* across $<w_@, w_1, \ldots, w_k>$, at a rate of 0.001mm per world. By the end of the sequence, β has moved an entire metre away. Throughout the series, Avril's perception of β is held fixed. It is determinate that Avril perceives β accurately in $w_@$. It is determinate that Avril does not perceive β accurately in w_k. But for many worlds w_i, positioned near the middle of the Sorites sequence, it is *vague* whether Avril perceives β accurately in w_i.

Henceforth, let us adopt the assumption that we occupy one of the worlds w_i in which it is vague whether Avril perceives β accurately. I pose the following question: *what follows from that fact that it is vague whether Avril perceives β accurately?*

On one view, it follows that there is some vague property *V* which Avril determinately predicates of β. It is vague whether Avril perceives β accurately because it is vague whether β instantiates *V*. Of course, this suggestion is wholly

[34] For the avoidance of doubt, I stress that this assertion is compatible with any sensible theory of vagueness, including epistemicism (Williamson 1994) and other classical-logic-endorsing views.

antithetical to precise supervaluationism. There are no vague properties in the precise supervaluationist's ontology. How, then, should the proponent of precise supervaluationism explain the fact that it is vague whether Avril perceives β accurately?

To answer this question, let us return to the simpler case of *vague predicates* such as '*x* is a small number', discussed above. According to precise supervaluationism, various precise properties are small-candidates; it is vague which of them is expressed by the predicate '*x* is a small number'. Furthermore, it is vague whether the latter predicate applies to the number 20 because 20 instantiates some (but not all) small-candidates. This account can be straightforwardly extended to the case of *vague perceptual accuracy*, along the following lines:

- Say that a property is 'relevant' iff it is a colour property, or a conjunctive property including some colour property as a conjunct.
- There is no relevant property *V* such that Avril determinately predicates *V*—just as there is no particular property *V* such that '*x* is a small number' determinately expresses *V*.
  Formally: $\sim\exists x(\text{relevant}(x)\ \&\ \Delta(\text{Avril predicates } x))$.
- Rather, there are various colour properties in the vicinity, and it is vague which of them Avril predicates—just as there are various precise arithmetical properties in the vicinity, and it is vague which of them is expressed by '*x* is a small number'.
  Formally: $\exists y_1 \ldots \exists y_n(\nabla(\text{Avril predicates } y_1)\ \&\ \ldots\ \nabla(\text{Avril predicates } y_n))$.
- If it is vague whether Avril predicates *y*, then *y* is a 'content-candidate'. The content-candidates comprise certain precise colour properties (corresponding to precise regions of colour space),[35] and conjunctive properties containing such colours as conjuncts. Picture the matter thus:

[35] It is important to stress that content-candidates need not be *superdeterminate* colours, in the sense of CHAPTER ONE §**3**. Content-candidates may also be *determinable* colours. The crucial point is that each content-candidate *c* is precise; there is no object *x* such that it is vague whether *x* instantiates *c*. The notion of a 'precise determinable property' is perfectly coherent.

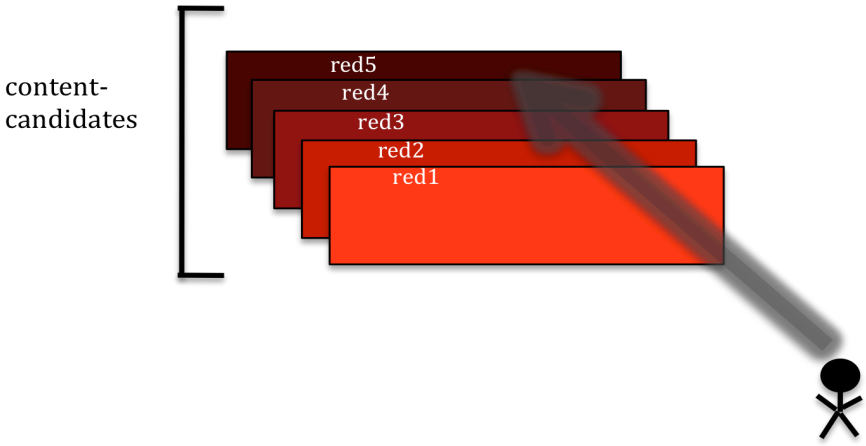

Figure 4: the coloured rectangles represent various content-candidates; it is vague which of them Avril predicates of β.

- It is vague whether '*x* is a small number' applies to the number 20 because 20 instantiates some (but not all) small-candidates. Likewise, it is vague whether Avril perceives β *accurately* because β instantiates some (but not all) content-candidates.

It seems to me that precise supervaluationists are required to adopt this account. In any case, I cannot readily envisage any alternatives. If I am right, then every precise supervaluationist is committed to assumption (A):

A. *Perceptual predication is vague*

There is no relevant property which Avril determinately predicates.

$\sim\exists x(\text{relevant}(x)\ \&\ \Delta(\text{Avril predicates } x))$

That is the first implication of precise supervaluationism on the metaphysics of perception, and it is by no means a trivial one.

Precise supervaluationism has a second important implication:

B. *Phenomenal characters are precise*

For all phenomenal characters *x*: it is not vague whether Avril instantiates *x*.

$\forall x(\text{phenomenal}(x) \rightarrow \sim\nabla(\text{Avril instantiates } x))$

After all, precise supervaluationists accept the following universal generalization: $\forall x \forall y \sim \nabla(y$ instantiates $x)$. As noted above, 'Avril' is a free variable, standing for an arbitrary perceiver. Therefore, it is permissible to instantiate 'Avril' in the latter generalization to obtain: $\forall x \sim \nabla$(Avril instantiates x). This trivially implies (B).[36]

§4 The Proof

(A) and (B) generate serious problems for representationalism. The difficulty should be pre-theoretically apparent: representationalism claims that there is an extremely close connection between phenomenal character and perceptual predication; (A) and (B) suggest otherwise. According to (A), perceptual predication is a vague matter; according to (B), phenomenal character is not. Indeed, provided that we make one prima facie plausible assumption, (A) and (B) are provably inconsistent with representationalism.

Consider the following claims:

[36] Why have I continually reminded the reader that 'Avril' is a free variable standing for an arbitrary object? It is widely recognized that the following inference is valid, where α is a free variable standing for an arbitrary object:

$\forall x \Delta \phi(x)$
$\therefore \Delta \phi(\alpha)$

However, it is equally widely recognized that same inference is *invalid* if α is a *vague noun phrase* (cf. Lewis 1988; Williamson 2003; Williams 2008). It is important, therefore, to stress that 'Avril' is a variable standing for an arbitrary object, not a vague proper name. Similar remarks apply to inferences such as:

$\Delta \phi(\alpha)$
$\therefore \exists x \Delta \phi(x)$

$\nabla \phi(\alpha)$
$\therefore \exists x \nabla \phi(x)$

I will routinely employ such inferences, but *only* when α is a free variable standing for an arbitrary object. More on this point in CHAPTER THREE §**1**.

REPRESENTATIONALISM

$\forall x(\text{phenomenal}(x) \rightarrow \exists y \forall z(z \text{ instantiates } x \leftrightarrow z \text{ predicates } y)$

For every phenomenal character x, there is some property y such that a subject z instantiates x iff z predicates y.[37]

Δ-REPRESENTATIONALISM

$\forall x(\text{phenomenal}(x) \rightarrow \exists y \forall z \Delta(z \text{ instantiates } x \leftrightarrow z \text{ predicates } y)$

For every phenomenal character x, there is some property y such that for any subject z <u>it is determinate that</u>: z instantiates x iff z predicates y.

The 'plausible assumption' required to prove the incompatibility of representationalism with (A) and (B) is this: *if representationalism is true, then Δ-representationalism is also true*. In CHAPTER FOUR, we shall explore in more detail whether representationalists should countenance denying this assumption. But there is no doubt that the assumption is pre-theoretically attractive: I doubt that many card-carrying representationalists are happy to deny Δ-representationalism. (Personal correspondence has confirmed this impression.) Using the modal logic for Δ outlined in **§1**, together with basic classical logic, we shall prove that (A), (B) and Δ-representationalism collectively lead to contradiction.

Avril is conscious. Therefore, there is at least one phenomenal character which Avril instantiates. This follows from two claims: (*i*) phenomenal characters are superdeterminates of phenomenal consciousness; and (*ii*) nothing can have a determinable property without having some superdeterminate thereof. (See CHAPTER ONE for discussion.) Let $\prod$ be the most complex perceptual phenomenal character which Avril instantiates. Δ-representationalism entails that:

1. $\exists y \forall z \Delta(z \text{ instantiates } \prod \leftrightarrow z \text{ predicates } y)$

There is some property y such that for every subject z, it is determinate that: z instantiates $\prod$ iff z predicates y.

[37] Strictly speaking, representationalism is only a thesis about *perceptual* phenomenal characters. See CHAPTER ONE **§5**. We can ignore this complication for present purposes.

Let σ be an arbitrary property, and suppose for reductio that:

2. $\forall z \Delta(z \text{ instantiates } \prod \leftrightarrow z \text{ predicates } \sigma)$

For every subject *x*, it is determinate that: *x* instantiates ∏ iff *z* predicates σ.

Evidently, σ is a complex conjunctive property, containing a variety of properties as conjuncts. Furthermore, one of its conjuncts must surely be a colour property. Otherwise it would be manifestly false that every subject who predicates σ instantiates ∏. Thus, σ is *relevant* (in the sense characterized in **§3**):

3. relevant(σ)

Since 'Avril' is a variable standing for an arbitrary perceiver, we can plug 'Avril' into (2) to obtain:

4. $\Delta(\text{Avril instantiates } \prod \leftrightarrow \text{Avril predicates } \sigma)$

By assumption,

5. Avril instantiates ∏

(B) entails:

6. $\sim\nabla(\text{Avril instantiates } \prod)$

In the weak modal logic outlined in **§1** containing the axiom (T_Δ), (5) and (6) entail:

7. $\Delta(\text{Avril instantiates } \prod)$

Applying the (K_Δ)-schema, (4) and (7) entail:

8. $\Delta(\text{Avril predicates } \sigma)$

(8) and (3) entail:

9. relevant(σ) & Δ(Avril predicates σ)

Since σ is a variable standing for an arbitrary property, we can employ ∃-introduction to obtain:

10. $\exists x$(relevant(x) & Δ(Avril predicates x))

According to (10), there is some property—presumably a very complex property—which Avril determinately predicates. But this straightforwardly contradicts (A). Therefore, by *reductio*, Δ-representationalism is false. If representationalism entails Δ-representationalism, then representationalism is also false.

Before we move on, let us pause to comment on two features of the argument developed above. First: Δ-representationalism is a biconditional claim. As such, it can be split into two components:

LEFT-TO-RIGHT

$\forall x$(phenomenal(x) → $\exists y \forall z$Δ(z instantiates x → z predicates y)

RIGHT-TO-LEFT

$\forall x$(phenomenal(x) → $\exists y \forall z$Δ(z predicates y → z instantiates x)

The argument advanced above only targets the left-to-right direction of Δ-representationalism. As inspection of the argument will reveal, it is the left-to-right direction which enables us to derive (4), leading inexorably to the conclusion above. I have said nothing to challenge the right-to-left direction of Δ-representationalism. One reaction to the argument, therefore, is to abandon the left-to-right direction of Δ-representationalism, but to preserve the right-to-left direction. I will assume, however, that representationalists would prefer to hold onto the full-strength biconditional.

Second: one of the most common strategies for arguing against representationalism involves playing the *counterexample game*. The basic goal of the game is to produce a plausible case in which two subjects instantiate the same

phenomenal character, but predicate different properties (or vice versa). The argument developed in this chapter is of an entirely different breed. It purports to provide a *nonconstructive* proof that representationalism is false. No concrete counterexample to representationalism has been produced; we have merely argued that representationalism is false on abstract and general grounds pertaining to the nature of vagueness, together with relatively uncontroversial logic.

Vague Properties

CHAPTER THREE

How should representationalists respond to the problem of vagueness? That is the central question discussed in the following two chapters.

There are only two serious options available (apart from embracing the problem of vagueness, and dispensing with representationalism altogether):

R1. Deny precise supervaluationism.

R2. Accept representationalism, and deny Δ-representationalism.

The present chapter explores response (R1); the following chapter explores response (R2). Let us begin with a brief recap. According to precise supervaluationism,

PRECISE SUPERVALUATIONISM

- Every property is precise.
- For any monadic predicate *F* and any object *x*: it is vague whether *F* applies to *x* iff *x* instantiates some (but not all) *F*-candidates.

If there is anything objectionable here, the source of the problem presumably lies in the claim that every property is precise. (A property *y* is precise iff $\Box{\sim}\exists x\nabla(x \text{ instantiates } y)$; otherwise *y* is a vague property.) The second claim—that it is vague whether a predicate *F* applies to an object *x* iff *x* instantiates some (but not all) *F*-candidates—seems virtually inevitable once vague properties are banned from ontology. There is room in logical space for intermediate views,[38] but I shall assume that the primary motivation for rejecting precise supervaluationism is the conviction that vague properties exist.

If vague properties exist, and precise supervaluationism is false, then the problem of vagueness is no longer a problem at all. Without precise supervaluationism, we no longer have any reason to accept premise (A)—the claim

[38] See Dietz and Moruzzi (2010) for a selection of such 'intermediate views'.

that there is no relevant property which Avril determinately predicates. To the contrary: if vague properties exist, then surely there *is* a relevant property *V* which Avril determinately predicates. *V* is a vague property; it is vague whether Avril perceives β accurately because it is vague whether β instantiates *V*. Likewise, there is no longer any reason to accept premise (B)—the claim that it is never vague whether a subject instantiates a given phenomenal character. This premise was motivated entirely by the view that vague properties do not exist.

Unfortunately, there are many good reasons to deny that vague properties exist. I shall argue that all philosophers—representationalists included—should strive to avoid commitment to vague properties. The problem of vagueness cannot be defused by admitting vague properties into one's ontology.

This chapter proceeds as follows. §**1** introduces the most obvious argument for believing in vague properties: the 'incredulous stare' argument. As we shall see, the incredulous stare argument is an unmitigated failure. Subsequently, §**2** develops two positive arguments against vague properties.

§1 The Incredulous Stare

§1.1 Warm-Up

Perhaps the most comprehensive recent discussion of vague properties is due to Williamson (2003).[39] We shall draw inspiration from his discussion at several points in this chapter—though, unlike Williamson, our discussion is couched in a first-order language instead of a higher-order language.

We shall assume that proponents of vague properties accept each of the following claims:

i. $\exists x \exists y \nabla$(*x* instantiates *y*)

There is some object and some property such that it is vague whether the former instantiates the latter.

ii. $\exists x \Delta$('small' expresses *x*)

There is some property *x* such that the adjective 'small' determinately expresses *x*.

[39] See also Schiffer (2003).

iii. $\exists x \Delta(\text{smallness} = x)$

There is some property x such that smallness is determinately identical to x.

As usual, 'small' and 'smallness' express the property of being a small number.

Given plausible assumptions, (*i*)—(*iii*) are equivalent. (*i*) leads inexorably to (*ii*) and (*iii*): if there are any vague properties, surely one of them is determinately expressed by the word 'small'. Likewise, if there are any vague properties, surely one of them is determinately identical to smallness. Conversely, both (*ii*) and (*iii*) lead inexorably to (*i*): if some property x is determinately expressed by 'small', then evidently there are various numbers n such that it is vague whether n instantiates x.

Of course, proponents of precise supervaluationism deny (*i*), (*ii*) and (*iii*). It is important to stress, however, that all mainstream precise supervaluationists accept that 'small' expresses *some* property, and that *some property* is identical to smallness (cf. Fine 1975; McGee and McLaughlin 1995; Williamson 2003). This is a simple observation, but it is often ignored by proponents of vague properties (e.g. Tye 1990). Precise supervaluationists merely insist that it is vague *which* property 'small' expresses, and that it is vague *which* property is identical to smallness. In other words: there is no property which 'small' *determinately* expresses, and no property to which smallness is *determinately* identical. As Braun and Sider (2007) observe, it does not follow that we should stop asserting sentences like 'the number 0 instantiates smallness'. Rather, we should simply recognize that it is not wholly clear which property the latter sentence ascribes to the number 0.

Is there any reason to believe in vague properties? I suspect that many philosophers who believe in vague properties simply find the following line of reasoning pre-theoretically compelling:

1. It is vague whether 'small' applies to the number 20.
2. Therefore, it is vague whether 20 is small.
3. Therefore, it is vague whether smallness is instantiated by 20.
4. Therefore, there is some property such that it is vague whether it is instantiated by 20.

5. Therefore, there is some object and some property such that it is vague whether the former instantiates the latter.

Opponents of vague properties accept (1), (2), and even (3). These philosophers claim that (1), (2) and (3) are all true for the same reason, viz. that the number 20 instantiates some (but not all) small-candidates. This is part and parcel of precise supervaluationism. But opponents of vague properties reject (4), on the grounds that every small-candidate is fully precise.

At this stage, proponents of vague properties typically adopt an incredulous stare. *How* on *earth* could *anyone* deny that (3) entails (4)? Call this the 'incredulous stare' argument. My goal in this section is to suggest that the incredulous stare argument is very far from decisive. In this section, I will not attempt to provide positive arguments for the claim that vague properties do not exist—merely to undermine the most obvious reason for thinking that they do.[40]

The move from (3) to (4) relies on the following inference:

INFERENCE A

$\nabla\phi(\sigma)$

$\therefore \exists x \nabla\phi(x)$

σ is a singular term.[41] $\phi(\sigma)$ is any well-formed formula *f* containing σ; $\phi(x)$ is the result of replacing at least once occurrence of σ in *f* with the variable *x* in *f*. Informally, inference A licenses reasoning such as: 'it is vague whether smallness is instantiated by the number 20; therefore, there is some property such that it is vague whether it is instantiated by the number 20.'

It may appear obvious that inference A is truth-preserving. But appearances can be deceptive. Let us begin with two 'warm-up' observations. First: there are many operators θ such that existential quantification into noun-phrase position within the scope of θ is not truth-preserving. Consider the operator 'it is contingent whether':

[40] To my knowledge, recent literature contains two other arguments for vague properties. The first argument, due to Schiffer (2003), appeals to intuitions about propositional attitudes; the second argument appeals to 'fuzzy logic' (see Williamson 2003 for discussion). There is no space to explore these arguments here.
[41] I will allow context to disambiguate use/mention distinctions.

6. It is contingent whether the number of planets = 9.
7. Therefore, there is some x such that it is contingent whether $x = 9$.

The premise of this argument is obviously true: 17 might have been the number of planets. But the conclusion of this argument is obviously false: the number 9 could not have been identical to 17; contingent identity is impossible (Kripke 1980). Might the inference from (3) to (4), and other related instances of inference (A), be equally dubious?

Here is one way of cashing out the analogy (Lewis 1988; Williamson 2003; Williams 2008). 'The number of planets' is a *nonrigid designator*: it denotes different numbers with respect to different possible worlds. Call the objects that 'the number of planets' denotes with respect to different possible worlds *number-contenders*. (6) effectively says: some (but not all) number-contenders are identical to the number 9. Obviously, this doesn't entail that the number 9 could have been identical to anything other than itself. Likewise, according to precise supervaluationism, 'smallness' is an *indeterminate designator*: it is vague which property it denotes. There are, in other words, various *small-candidates*. (3) effectively says: the number 20 instantiates some (but not all) small-candidates. Why should this entail the existence of a property such that it is vague whether the number 20 instantiates it?

Here is a second 'warm-up' observation. Say that n is the *last small number* iff n is small, and n+1 is not small. Consider the following inference:

8. It is determinate that (the last small number = the last small number).
9. Therefore, there is some number n such that it is determinate that (n = the last small number).

The premise of this argument is obviously true. Yet the conclusion is obviously false: in effect, (9) states that the boundary between small numbers and nonsmall numbers is a precise one, and this is plainly not so. Since (8) does not entail (9), we should be extremely suspicious of inference A. To be sure, the move from (8) to (9) involves quantifying into description-position within the scope of the operator Δ, whereas inference A only licenses quantification into singular-term position within the scope

of the operator ∇. Nevertheless, such examples vividly illustrate the difficulties associated with 'quantifying in' to noun-phrase position within the scope of operators pertaining to vagueness and determinacy.

§1.2 An Argument Against Inference A

So much for the warm-up. I shall now develop a positive argument against the validity of inference A. (For our purposes, to say that inference A is invalid is to say that not every instance of inference A is truth-preserving.) The argument does not rely on any aspect of precise supervaluationism. Since inference A is invalid, the 'incredulous stare' argument for vague properties relies on an invalid form of reasoning, and carries no persuasive force.

My argument against inference A relies on a famous theorem, due to Gareth Evans (1978), to the effect that identity is precise.

EVANS' THEOREM

$\forall x \forall y \sim \nabla(x = y)$

For any x and any y, it is not vague whether $x = y$.

Evans' argument for the precision of identity is extremely elegant. The proof appeals to a schematic version of Leibniz's law:

$\forall x \forall y[x = y \rightarrow (\phi(x) \leftrightarrow \phi(y)]$

For any x and y: if $x = y$, then $\phi(x)$ iff $\phi(y)$.

As it stands, Leibniz's law is merely a schema. To obtain instances of the schema, $\phi(x)$ and $\phi(y)$ must be replaced with any two well-formed formulae f and f_+ which differ only in the following respect: the variable x occupies one or more of the same syntactic positions in f which the variable y occupies in f_+.[42] Leibniz's law is among the most fundamental facts about identity. Its basic motivation is simple. Every variable assignment which makes the open sentence '$x = y$' true must assign the

[42] Furthermore, the variable x must not be bound by any quantifier which occurs within f, and the variable y must not be bound by any quantifier which occurs within f_+—though these variables may be bound by quantifiers which do not occur within f or f_+. See van Dalen (2008) for technical details.

variables x and y to exactly the same object. Thus, such assignments will always determine the same truth-value for any well-formed formulae $\phi(x)$ and $\phi(y)$.

Let w and z be arbitrary objects, and suppose for *reductio* that it is vague whether $w = z$. Following Evans, we argue as follows:

1. $\Delta(w = w)$ [Everything is determinately identical to itself, after all!]
2. $\sim\Delta(w = z)$ [By hypothesis, it is vague whether $w = z$]

But by Leibiniz's law,

3. $w = v \rightarrow (\Delta(w = w) \leftrightarrow \Delta(w = v)]$

By contraposition, (1), (2) and (3) entail:

4. $w \neq z$

Generalizing, for any x and y: if it is vague whether $x = y$, then $x \neq y$. Given plausible assumptions—relegated to a footnote—it follows that for any x and y, it is not vague whether $x = y$.[43]

Evans' argument was regarded as controversial in the years immediately following its publication. But the precision of identity is now accepted by the large majority of philosophers of vagueness, across a multitude of theoretical perspectives. At any rate, I shall assume that Evans' theorem is indeed a theorem. This assumption has dialectical force: proponents of vague properties typically concede that identity is precise (e.g. Tye 1990; Akiba 2000).

---

[43] If are any cases of vague identity, then surely there are also cases of *determinately* vague identity: $\Delta\nabla(w = z)$. On this assumption, we argue as follows (cf. Heck 1998):

i.	$\nabla(w = z) \rightarrow \sim(w = z)$	[proven above]
ii.	$\Delta(\nabla(w = z) \rightarrow \sim(w = z))$	[(N_Δ): *i*]
iii.	$\Delta\sim(w = z)$	[(K_Δ): *ii* + assumption]
iv.	$\sim\nabla(w = z)$	[definition of Δ: *iii*]

But (*iv*) contradicts our assumption that $\Delta\nabla(w = z)$. So there are no cases of determinately vague identity, and hence no cases of vague identity.

My central submission is straightforward*: inference A generates violations of Evans' theorem*. Here's why.

We begin by introducing the name 'Fuzzy', *via* the following reference-fixing description:

'Fuzzy' refers to the last small number.

The last small number is the number n such that n is small and n+1 is not small. (See §**1.1**.) It is important to be clear: 'the last small number' is *not* a nonreferring description, analogous to 'the King of France'. To the contrary, classical logic guarantees that *there is* a last small number; the existence of such a cut-off point is now widely accepted by philosophers of vagueness.[44] As such, it is incoherent to deny that Fuzzy exists. Of course, these remarks are compatible with the platitudinous observation that it is vague *which* number is the last small number. Among other things,

1. ∇(20 = the last small number)
   It is vague whether 20 = the last small number.

If you don't like the number 20, choose some other number that you prefer. Since smallness-talk is context-sensitive, there should be no difficulty in simply stipulating that we occupy a context in which (1) holds.

If (1) is true, then evidently (2) also holds:

2. ∇(Fuzzy = 20)
   It is vague whether Fuzzy = 20.

In light of the description we used to fix the reference of the name 'Fuzzy', (2) appears virtually inevitable. Applying inference A, it follows that:

3. $\exists x \nabla(x = 20)$

[44] *Proof*: if $\sim\exists n$(n is the last small number) then $\sim\exists n$(n is small & n+1 is not small), which classically entails that $\forall n$(n is small $\rightarrow$ n+1 is small). Yet the latter claim is plainly absurd.

There is some x such that it is vague whether $x = 20$.

Applying inference A again, it follows that:

4. $\exists x \exists y \nabla(x = y)$

There is some x and some y such that it is vague whether $x = y$.

Yet (4) contradicts Evans' theorem that identity is precise. We have a contradiction. What has gone wrong?

The culprit, I submit, is inference A. We should obviously accept that it is vague whether Fuzzy = 20. Nevertheless, it does *not* follow that there is some number x such that it is vague whether $x = 20$. It is may help to view that matter from the perspective of precise supervaluationism. 'Fuzzy' is an imprecise name. As such, there are various 'Fuzzy-candidates': the number 18, the number 19, the number 20, and so forth. One such candidate is determinately identical to the number 20; every other such candidate is determinately distinct from the number 20. *That* is why it is vague whether Fuzzy = 20. It certainly does not follow, however, that there is some number x such that it is vague whether $x = 20$. Inference A is simply invalid.

Defenders of inference (A) may be tempted to pursue an alternative response, and deny that (1) entails (2). On this view,

1. It is vague whether 20 = the last small number.

is true, and

2. It is vague whether Fuzzy = 20.

is false. This position, however, is utterly unstable. In light of the description we employed to fix the reference of the name 'Fuzzy', the following biconditional is knowable *a priori*: (Fuzzy = 20 iff 20 = the last small number). Every proposition which is knowable *a priori* is determinately true. It follows, therefore, that:

5. Δ(Fuzzy = 20 iff 20 = the last small number).

In CHAPTER TWO §**1**, we proved that the following 'V-schema' holds in a weak modal logic: $\Delta(\phi \leftrightarrow \psi) \rightarrow (\nabla\phi \leftrightarrow \nabla\psi)$. Since it is vague whether 20 = the last small number, the V-schema entails that it is vague whether Fuzzy = 20. As such, (1) implies (2).

Conclusion: inference A is invalid.

Let us take stock. We began by confronting the 'incredulous stare' argument for vague properties, which insists that the following inference is obviously truth-preserving:

It is vague whether 20 instantiates smallness.
Therefore, there is some property such that it is vague whether 20 instantiates it.

We observed that such reasoning is an instance of a more general inference rule:

INFERENCE A

$\nabla\phi(\sigma)$

$\therefore \exists x \nabla\phi(x)$

I have developed an independent argument against the validity of inference A which does not appeal to any feature of precise supervaluationism. My argument reveals that inference A generates violations of Evans' theorem, which is unacceptable. The incredulous stare argument therefore relies on an invalid form of reasoning; it provides no convincing motivation for believing in vague properties.

(It is worth stressing that inference A is perfectly acceptable when σ is a *free variable* standing for an arbitrary object, introduced in the course of a proof using $\forall$ or $\exists$ introduction. Problems only arise when σ is a vague noun phrase like 'smallness'. See CHAPTER TWO note 36 for more on this point.)

§2 Two Positive Arguments Against Vague Properties

Inference A is invalid. It does not follow, of course, that vague properties do not exist. It is perfectly consistent to hold that at least some vague properties exist, while

admitting that certain instances of inference A are not truth-preserving. Although the invalidity of inference A undercuts the 'incredulous stare' argument, it is epistemically possible that alternative considerations favour the existence of vague properties. In order to provide positive reasons for denying the existence of vague properties, more work is required.

In this vein, the remainder of the present chapter develops two positive arguments against the existence of vague properties. §**2.1** formulates *the argument from haecceity*; §**2.2** formulates *the argument from extensionality*. I shall continue to assume that Evans' theorem holds, though I shall not assume that proponents of vague properties accept every instance of inference A. Taken together, the arguments developed below provide strong positive reason to deny that vague properties exist.

§2.1 The First Argument: Vagueness and Haecceities

The term 'haecceity'—a rather barbarous medieval Latin locution—has been recruited to serve a variety of functions in philosophy. For our purposes, it is convenient to define the term 'haecceity' by the following schema:

The property of *being identical to A* is the haecceity of *A*.

Thus, the property of being identical to Tim is the haecceity of Tim; the property of being identical to Avril is the haecceity of Avril. Haecceities give rise to a number of interesting metaphysical questions.[45] I shall argue that haecceities also generate serious problems for proponents of vague properties.

As we observed in §**1.1**, proponents of vague properties claim that:

1. $\exists x \Delta(x = \text{smallness})$

   There is some property x such that x is determinately identical to smallness.

Smallness is not a haecceity.[46] Nevertheless, proponents of vague properties presumably accept analogous claims about haecceities. For example, consider Fuzzy. (As stated in **§1**, Fuzzy is the last small number). Let H_{fuzzy} be the haecceity of Fuzzy.

[45] See Adams (1981) and Plantinga (1974) for discussion.
[46] Given plausible assumptions about the metaphysics of properties.

In other words, H_{fuzzy} is the property of being identical to Fuzzy. Presumably, proponents of vague properties hold that:

2. $\exists x \Delta(x = H_{\text{fuzzy}})$

There is some property x such that x is determinately identical to H_{fuzzy}.

Philosophers who accept (1) should surely accept (2); there is no obvious motivation for accepting (1) and denying (2). I shall argue, however, that (2) leads immediately to contradiction. My argument does not assume that proponents of vague properties accept inference A.

Begin with the following thought. *Each haecceity is individuated by the object it is 'about'*. More carefully, the following schema holds: (the property of being identical to A = the property of being identical to B) iff ($A = B$). In the jargon introduced above, (the haecceity of A = the haecceity of B) iff ($A = B$).

Certain hardcore Fregeans may be tempted to reject this schema. Perhaps such philosophers will insist that the property of being identical to Hesperus is distinct from the property of being identical to Phosphorus, even though Hesperus = Phosphorus. In practice, however, most Fregeans are careful to avoid any such claim. Fregeans typically distinguish properties from modes of presentation. According to a standard contemporary version of Fregeanism, expressions like 'the haecceity of Hesperus' and 'the haecceity of Phosphorus' are associated with different modes of presentation, but nevertheless denote the same property.

In any case, we shall assume that each haecceity is individuated by the object it is about, in the sense elucidated above. (Certainly, this principle is entailed by Russellian and set-theoretic accounts of properties.) It follows that the haecceity of Fuzzy and the haecceity of the number 20 are identical iff Fuzzy is identical 20. More carefully, where H_{20} is the haecceity of the number 20:

1. $\forall x \forall y((x = H_{\text{fuzzy}} \ \& \ y = H_{20}) \rightarrow (x = y \leftrightarrow \text{Fuzzy} = 20))$

For any x and any y: if $x = H_{\text{fuzzy}}$ and $y = H_{20}$, then ($x = y$ iff Fuzzy = 20).

Presumably, this principle is fully determinate:

2. $\forall x \forall y \Delta((x = H_{\text{fuzzy}} \ \& \ y = H_{20}) \rightarrow (x = y \leftrightarrow \text{Fuzzy} = 20))$

For any x and any y, it is determinate that: if $x = H_{\text{fuzzy}}$ and $y = H_{20}$, then ($x = y$ iff Fuzzy = 20).

In **§1**, I argued that:

3. $\nabla(\text{Fuzzy} = 20)$

It is vague whether Fuzzy = 20.

We now have a problem. Proponents of vague properties claim that some property is determinately identical to H_{fuzzy}, and that some property is determinately identical to H_{20}. That is:

4. $\exists x \Delta(x = H_{\text{fuzzy}})$

There is some property x such that x is determinately identical to H_{fuzzy}.

5. $\exists x \Delta(x = H_{20})$

There is some property x such that x is determinately identical to H_{20}.

Together, premises (2)—(5) are *extremely problematic*. In a classical weak modal logic for Δ, (2)—(5) generate a violation of Evans' theorem:

6. $\exists x \exists y \nabla(x = y)$

There is some x and some y such that it is vague whether $x = y$.

The proof that (2)—(5) entail (6) is provided in Figure 5, below. (Optional reading!) Since (6) contradicts Evans' theorem, *something must go*.

I submit that premise (4) is the culprit. Every other premise is unimpeachable. According to premise (4), something is determinately identical to H_{fuzzy}. If we dispense with this premise, it is no longer possible to derive a contradiction of Evans' theorem. However, dispensing with premise (4) *also* requires us to abandon the view that vague properties exist. For—as we have seen—the view that vague properties exist wrongly predicts that premise (4) is true.

Conclusion: vague properties do not exist.

Figure 5: The Proof

Let α be an arbitrary object such that $\Delta(\alpha = H_{\text{fuzzy}})$; let β be an arbitrary object such that $\Delta(\beta = H_{20})$. By (2),
$\Delta((\alpha = H_{\text{fuzzy}} \;\&\; \beta = H_{20}) \rightarrow (\alpha = \beta \leftrightarrow \text{Fuzzy} = 20))$. By (K_{Δ}),
$\Delta(\alpha = \beta \leftrightarrow \text{Fuzzy} = 20))$. The following 'V-schema' is a theorem of weak modal logic: $\Delta(\phi \leftrightarrow \psi) \rightarrow (\nabla\phi \leftrightarrow \nabla\psi)$. By (3), $\nabla(\text{Fuzzy} = 20)$. By the V-schema, $\nabla(\alpha = \beta)$. Since α and β are free variables standing for arbitrary objects, we may employ $\exists$-intro to obtain: $\exists x \exists y \nabla(x = y)$. Q.E.D.

How should proponents of vague properties respond to the foregoing argument? The best response, it seems to me, involves a defensive manoeuvre. Proponents of vague properties should *concede* that premise (4) is false: nothing is determinately identical to H_{fuzzy}. Nevertheless, such philosophers should stress that the falsehood of (4) is not—*strictly speaking*—inconsistent with their position. As we have seen, proponents of vague properties accept the following claim about smallness:

7. $\exists x \Delta(x = \text{smallness})$
   There is some property x such that x is determinately identical to smallness.

I suggested that proponents of (7) should also accept (4):

4. $\exists x \Delta(x = H_{\text{fuzzy}})$
   There is some property x such that x is determinately identical to H_{fuzzy}.

Perhaps, however, we moved too quickly. What if proponents of vague properties accept (7), but deny (4)? In effect, philosophers who accept (7) and deny (4) concede that there are no *vague haecceities* like H_{fuzzy}, but insist that there are *vague qualities* like smallness. Call this view 'the Halfway House'. Does the Halfway House stand any prospect of success?

§2.2 The Second Argument: Vagueness and Extensionality

This section develops an argument against the 'Halfway House' position introduced above. In other words, I shall argue that no property is determinately identical to smallness: $\sim\exists x\Delta(x = \text{smallness})$. If the argument succeeds, it is high time to dispense with the view that vague properties exist.[47]

For heuristic purposes, it will prove fruitful to begin with a simple version of the argument. On one view, properties are extensional: a property x is identical to a property y iff x and y are coextensional—that is, iff exactly the same objects instantiate x and y. Call this 'the principle of extensionality'. The principle of extensionality will appeal to philosophers who identify the property of being F with the set of Fs, or some related set-theoretic construction (see Lewis 1986). The principle of extensionality is controversial, but let us assume that it is true.

Let the variables x and y range over properties, and let $\boldsymbol{C}(x, y)$ hold iff x and y are coextensional. As usual, smallness is the property of being a small number; *eicosity* is the property of being a number less than 20. (The source of this term will be familiar to readers conversant with Greek.) According to the principle of extensionality,

1. $\forall x\forall y(x = \text{smallness} \;\&\; y = \text{eicosity} \rightarrow (x = y \leftrightarrow \boldsymbol{C}(\text{smallness}, \text{eicosity}))$
   For any property x and any property y: if x = smallness and y = eicosity, then $x = y$ iff smallness is coextensional with eicosity.

Presumably, (1) is fully determinate:

2. $\forall x\forall y\Delta(x = \text{smallness} \;\&\; y = \text{eicosity} \rightarrow (x = y \leftrightarrow \boldsymbol{C}(\text{smallness}, \text{eicosity}))$
   For any property x and any property y, it is determinate that: if x = smallness and y = eicosity, then $x = y$ iff smallness is coextensional with eicosity.

As we observed in §**1.2**,

[47] The argument in this section is indirectly inspired by the mereological arguments in Weatherson (2003). For discussion, see Barnes and Williams (2009).

3. $\nabla\forall x(x$ is small iff $x < 20)$

It is vague whether (a number is small iff it is less than 20).

If (3) is true, then (4) follows immediately:

4. $\nabla \boldsymbol{C}$(smallness, eicosity)

It is vague whether smallness and eicosity are coextensional.

We now have a problem. Proponents of vague properties accept that:

5. $\sim\exists x\Delta(x =$ smallness$)$

Some property is determinately identical to smallness.

6. $\sim\exists x\Delta(x =$ eicosity$)$

Some property is determinately identical to eicosity.

Taken together, premises (2)—(6) are *extremely problematic*. In a classical weak modal logic for Δ, (2)—(6) generate a violation of Evans' theorem:

7. $\exists x\exists y\nabla(x = y)$

There is some x and some y such that it is vague whether $x = y$.

The proof that (2)—(6) and the principle of extensionality entail (7) is provided in Figure 6, below. Since (7) contradicts Evans' theorem, *something must go*.

Figure 6: The Proof

Let α be an arbitrary property such that $\Delta(\alpha =$ smallness$)$; let β be an arbitrary property such that $\Delta(\beta =$ eicosity$)$. By (2), $\Delta((\alpha =$ smallness $\&\ \beta =$ eicosity$) \rightarrow (\alpha = \beta \leftrightarrow \boldsymbol{C}$(smallness, eicosity)$))$. By (K_Δ), $\Delta(\alpha = \beta \leftrightarrow \boldsymbol{C}$(smallness, eicosity)$)$. The following 'V-schema' is a theorem of weak modal logic: $\Delta(\phi \leftrightarrow \psi) \rightarrow (\nabla\phi \leftrightarrow \nabla\psi)$. By (3), $\nabla\boldsymbol{C}$(smallness, eicosity). Since α and β are free variables standing for arbitrary objects, we may employ $\exists$-intro to obtain: $\exists x\exists y\nabla(x = y)$. Q.E.D.

Those who subscribe to the principle of extensionality will evidently insist that premise (5) is the culprit. On this view, no property is determinately identical to smallness. Although I believe that this is the correct reaction, there is no doubt that proponents of vague properties will instead deny the principle of extensionality, which enabled us to obtain premise (2). After all, the principle of extensionality is extremely controversial, and faces a number of apparent counterexamples. Imagine, for example, that exactly the same objects instantiate the property of having a heart and the property of having a kidney. (As it happens, this empirical generalization is false—but it is illustrative.) The principle of extensionality entails that the property of having a heart and the property of having a kidney are thereby identical. Many are reluctant to accept this conclusion.

Fortunately, it is possible to formulate a version of the argument developed above which does not rely on the principle of extensionality. That is the task to which we now turn.

According to the principle of *modal extensionality*, modally equivalent properties are identical. That is: if x and y are properties which are instantiated by the same objects *in every possible world*, then $x = y$. Formally, where the variables x and y range over properties, this amounts to:

MODAL EXTENSIONALITY

$\forall x \forall y(x = y \leftrightarrow \Box \boldsymbol{C}(x, y))$

For any properties x and y: $x = y$ iff x and y are necessarily coextensional.

For obvious reasons, the principle of modal extensionality enjoys a far better reputation than the principle of extensionality *simpliciter* (discussed above). Many philosophers are suspicious of hyperintensional metaphysics; among such philosophers, the principle of modal extensionality has a strong following. This is not to say, of course, that the relevant principle is universally accepted. Opponents of modal extensionality insist that the property of *being red* and the property of *being red and such that 2+2=4* are not identical, even though these properties are necessarily coextensional. Proponents of modal extensionality deny this 'intuition', and hold that every other approach to individuating properties leads to abhorrent

cardinality problems (see Hawthorne and Uzquiano 2012 for a sample). I do not propose to engage with this controversy in any serious way. My argument will simply employ modal extensionality as a premise. The argument will not persuade philosophers who deny modal extensionality; nor will it purport to.

We proceed in a familiar manner. By modal extensionality,

1. $\forall x \forall y(x = \text{smallness} \ \& \ y = \text{eicosity} \rightarrow (x = y \leftrightarrow \Box \boldsymbol{C}(\text{smallness}, \text{eicosity}))$

   For any property x and any property y: if x = smallness and y = eicosity, then $x = y$ iff smallness is necessarily coextensional with eicosity.

Presumably, (1) is fully determinate:

2. $\forall x \forall y \Delta(x = \text{smallness} \ \& \ y = \text{eicosity} \rightarrow (x = y \leftrightarrow \Box \boldsymbol{C}(\text{smallness}, \text{eicosity}))$

   For any property x and any property y, it is determinate that: if x = smallness and y = eicosity, then $x = y$ iff smallness is necessarily coextensional with eicosity.

But once we have come this far, the game is over: if vague properties exist, (1) leads inexorably to violations of Evans' theorem.

As we observed above, it is vague whether smallness and eicosity are coextensional. Moreover, this fact is not a mere contingency: in *every* possible world, it is vague whether smallness and eicosity are coextensional. That is:

3. $\Box \nabla \boldsymbol{C}(\text{smallness}, \text{eicosity})$

   It is necessarily vague whether smallness and eicosity are coextensional.

It follows that:

4. $\nabla \Box \boldsymbol{C}(\text{smallness}, \text{eicosity})$

   It is vague whether it is necessary that smallness and eicosity are coextensional.

The argument from (3) to (4) is an instance of a plausible principle governing the interaction of necessity and modality: $\Box\nabla\phi \rightarrow \nabla\Box\phi$. (If it is necessarily vague whether ϕ, then it is vague whether it is necessary that ϕ.)

We now have a problem. Proponents of vague properties accept that:

5. $\sim\exists x\Delta(x = \text{smallness})$

   Some property is determinately identical to smallness.

6. $\sim\exists x\Delta(x = \text{eicosity})$

   Some property is determinately identical to eicosity.

Taken together, premises (2)—(6) are *extremely problematic*. In a classical weak modal logic for Δ, (2)—(6) generate a violation of Evans' theorem:

7. $\exists x\exists y\nabla(x = y)$

   There is some x and some y such that it is vague whether $x = y$.

The proof that (2)—(6) entail (7) is exactly the same as the proof provided in Figure *n*, above. Since (7) contradicts Evans' theorem, *something must go*. I submit that the appropriate response is to deny (5): no property is determinately identical to smallness. If so, the view that vague properties exist stands refuted.

How should the defender of vague properties respond? Is there any other option? Let ϕ abbreviate: ***C***(smallness, eicosity). When moving from premise (3) to (4), I relied on the assumption that $\Box\nabla\phi$ entails $\nabla\Box\phi$. I offered no argument for this assumption, and merely described the putative entailment as 'plausible'. Fortunately, the relevant entailment can be formally proven in a weak modal logic for $\Box$ and ∇. The proof is provided in Figure 7, below; those who are already convinced of the relevant entailment are free to skip it.

Figure 7: The Proof ($\Box\nabla\phi$ entails $\nabla\Box\phi$)

- *Assumption*: $\Box\phi \vee \Box{\sim}\phi$.
- This assumption is false for most ϕ, but true for the particular ϕ under consideration: if smallness and eicosity are coextensional, then smallness and eicosity are necessarily coextensional; if smallness and eicosity are not coextensional, then smallness and eicosity are necessarily not coextensional.
- To establish that $\Box\nabla\phi$ entails $\nabla\Box\phi$, we must prove two things:

1. $\Box\nabla\phi \rightarrow {\sim}\Delta\Box\phi$
2. $\Box\nabla\phi \rightarrow {\sim}\Delta{\sim}\Box\phi$

Proof of (1):

a.	$\Box\phi \rightarrow \phi$	($T_\Box$)
b.	$\Delta(\Box\phi \rightarrow \phi)$	(N_Δ, *a*)
c.	$\Delta\Box\phi \rightarrow \Delta\phi$	(K_Δ, *b*)
d.	$\Delta\phi \rightarrow {\sim}\nabla\phi$	(definition of ∇, Δ)
e.	${\sim}\nabla\phi \rightarrow {\sim}\Box\nabla\phi$	($T_\Box$)
f.	$\Box\nabla\phi \rightarrow {\sim}\Delta\Box\phi$	(by *c*, *d*, *e*)

Proof of (2):

g.	${\sim}\Box\phi \rightarrow {\sim}\phi$	(by the 'assumption', above)
h.	$\Delta({\sim}\Box\phi \rightarrow {\sim}\phi)$	(N_Δ, *g*)
i	$\Delta{\sim}\Box\phi \rightarrow \Delta{\sim}\phi$	(K_Δ, *h*)
j.	$\Delta{\sim}\Box\phi \rightarrow {\sim}\nabla\phi$	(by *i* and the definition of Δ, ∇)
k.	${\sim}\nabla\phi \rightarrow {\sim}\Box\nabla\phi$	($T_\Box$)
l.	$\Delta{\sim}\Box\phi \rightarrow {\sim}\Box\nabla\phi$	(*j*, *k*)
m.	$\Box\nabla\phi \rightarrow {\sim}\Delta{\sim}\Box\phi$	(*l*)

The only remaining option available to defenders of vague properties is to deny the principle of modal extensionality. As noted above, it is not feasible to engage with the debate between proponents and opponents of modal extensionality in any serious way

here. Nevertheless, if defenders of vague properties are forced to abandon a highly attractive approach to the individuation of properties, this is a serious concern. I suspect that many philosophers will prefer an alternative conclusion: no property is determinately identical to smallness; vague properties do not exist.

…

At the beginning of this chapter, we observed that if vague properties exist, representationalism escapes the problem of vagueness developed in CHAPTER TWO. In response to this observation, I have developed two positive arguments against vague properties: the argument from haecceity, and the argument from extensionality. Both arguments show that vague properties generate violations of Evans' theorem, given plausible assumptions.

Conclusion: representationalists are well-advised to seek out an alternative solution to the problem of vagueness, which does not incur commitment to vague properties.

Δ-Representationalism

CHAPTER FOUR

In CHAPTER TWO, we distinguished two views:

REPRESENTATIONALISM

$\forall x(\text{phenomenal}(x) \rightarrow \exists y \forall z(z \text{ instantiates } x \leftrightarrow z \text{ predicates } y)$

For every phenomenal character x, there is some property y such that a subject z instantiates x iff z predicates y.

Δ-REPRESENTATIONALISM

$\forall x(\text{phenomenal}(x) \rightarrow \exists y \forall z \Delta(z \text{ instantiates } x \leftrightarrow z \text{ predicates } y)$

For every phenomenal character x, there is some property y such that for any subject z <u>it is determinate that</u>: z instantiates x iff z predicates y.

Thus far, we have assumed that proponents of representationalism should *also* accept Δ-representationalism. This assumption is crucial: without it, the 'problem of vagueness' developed in CHAPTER TWO simply cannot arise within a classical weak modal logic for Δ.

Might representationalists avoid the problem of vagueness by denying Δ-representationalism? That is the central question discussed in this chapter. It is certainly difficult to envisage a card-carrying representationalist happily asserting the negation of Δ-representationalism. But is there any *substantive* argument that anyone who accepts the former should accept the latter, or must we resort to table-pounding?

§1 Warm-Up

§1.1 Arguing by Analogy: □ and Δ

Here is a natural thought. If representationalism is true, then representationalism is *fully necessary*. In other words, we should be able to enrich representationalism with a necessity operator in the following positions, without changing its truth-value:

$\Box\forall x(\text{phenomenal}(x) \rightarrow \exists y\forall z(z \text{ instantiates } x \leftrightarrow z \text{ predicates } y)$

$\forall x(\text{phenomenal}(x) \rightarrow \Box\exists y\forall z(z \text{ instantiates } x \leftrightarrow z \text{ predicates } y)$

$\forall x(\text{phenomenal}(x) \rightarrow \exists y\Box\forall z(z \text{ instantiates } x \leftrightarrow z \text{ predicates } y)$

$\forall x(\text{phenomenal}(x) \rightarrow \exists y\forall z\Box(z \text{ instantiates } x \leftrightarrow z \text{ predicates } y)$

Representationalism is, after all, a bona fide metaphysical thesis. In general, the truth-value of a genuine metaphysical thesis should be preserved when it is enriched with necessity operators along the lines depicted above.

For similar reasons, the following conditional has apparent plausibility: if representationalism is true, then representationalism is *fully determinate*. In other words, if the truth-value of representationalism is preserved when it is enriched with a necessity operator in a certain location, its truth-value should also be preserved when it is enriched with a corresponding determinacy operator:

$\Delta\forall x(\text{phenomenal}(x) \rightarrow \exists y\forall z(z \text{ instantiates } x \leftrightarrow z \text{ predicates } y)$

$\forall x(\text{phenomenal}(x) \rightarrow \Delta\exists y\forall z(z \text{ instantiates } x \leftrightarrow z \text{ predicates } y)$

$\forall x(\text{phenomenal}(x) \rightarrow \exists y\Delta\forall z(z \text{ instantiates } x \leftrightarrow z \text{ predicates } y)$

$\forall x(\text{phenomenal}(x) \rightarrow \exists y\forall z\Delta(z \text{ instantiates } x \leftrightarrow z \text{ predicates } y)$

Representationalism is, after all, a bona fide metaphysical thesis. The truth-value of a genuine metaphysical thesis should, in general, be preserved when it is enriched with determinacy operators along the lines depicted above. Or so it seems. Since the final enrichment listed above is simply a restatement of Δ-representationalism, it follows that proponents of representationalism should accept Δ-representationalism.

But this argument moves too quickly. In certain cases, a 'bona fide metaphysical thesis' appears to be fully necessary, yet not fully determinate. Consider the following claims:

1. $\exists n\Box(n \text{ is small } \& \ n+1 \text{ is not small})$
2. $\exists n\Delta(n \text{ is small } \& \ n+1 \text{ is not small})$

(1) is widely accepted. As we observed in CHAPTER THREE §**1.2** (note 44), classical logic guarantees that $\exists n$(n is small & n+1 is not small). Presumably, the latter fact is fully necessary: $\exists n\Box$(n is small & n+1 is not small).

On the other hand, everyone denies (2). Why? Distinguish two claims:

2. $\exists n\Delta$(n is small & n+1 is not small)
   There is some particular number such that *it* is determinately the last small number.

3. $\Delta\exists n$(n is small & n+1 is not small)
   It is determinate that *some number or other* is the last small number.

It is widely agreed that (2) and (3) have different truth-conditions. As usual, a *small-candidate* is a property S such that it is vague whether 'small' expresses S. (See CHAPTER TWO §**2** for details). According to the standard supervaluationist semantics (e.g. Fine 1975; McGee and McLaughlin 1995),

- (2) is true iff there is a some number n such that for every small-candidate S, n instantiates S and n+1 does not instantiate S.
- (3) is true iff for every small-candidate S, there is a some number n such that n instantiates S and n+1 does not instantiate S.

The truth-condition for (2) has the form: 'there is some y such that for every x, …'
The truth-condition for (3) has the form: 'for every x, there is some y such that…'
As every philosopher knows, such scope distinctions are extremely important: compare $\exists y\forall x$(x loves y) and $\forall x\exists y$(x loves y).

So understood, (3) is true and (2) is false. To see that this is so, consider the following small-candidates:

- the property of being less than 18;
- the property of being less than 19;
- the property of being less than 20;
- the property of being less than 21.

Say that n is the *cut-off point for S* iff n instantiates S and n+1 does not.
17 is the cut-off point for the first property; 18 is the cut-off point for the second property; 19 is the cut-off point for the third property; 20 is the cut-off point for the fourth property. As these remarks indicate, each small-candidate has *a* cut-off point. (3) is therefore true. Yet each small-candidate has a *different* cut-off point. (2) is therefore false, for (2) requires that each small-candidate has the *same* cut-off point.

Given plausible assumptions, then, it is not always possible to simply remove the necessity operator □ and replace it with the determinacy operator Δ in the same position. Thus, it is wrong to automatically assume that (4) entails (5):

4. $\exists y \forall x \Box(x \text{ instantiates } \prod \leftrightarrow x \text{ predicates } y)$
5. $\exists y \forall x \Delta(x \text{ instantiates } \prod \leftrightarrow x \text{ predicates } y)$

§1.2 Delving Further: The Toy Model

Although it is *consistent* to accept (4) and deny (5), it is natural to wonder: what would the world have to be like, in order for representationalism to turn out true and Δ-representationalism to turn out false?

Here, so far as I can see, is the only possible answer. We must assume that no particular relation is determinately expressed by the central piece of representationalist ideology: 'z predicates y'. Instead, there are various relations in the vicinity, and it is vague which of them is expressed by 'z predicates y'. In familiar terminology, there are multiple *predication-candidates*: predication_1, predication_2, and so forth. (We shall assume, for simplicity, that there are exactly two predication-candidates, although nothing in our discussion will turn on this assumption.) When a subject perceives the world, she predicates_1 one property (call it F) and simultaneously predicates_2 a distinct property (call it G):

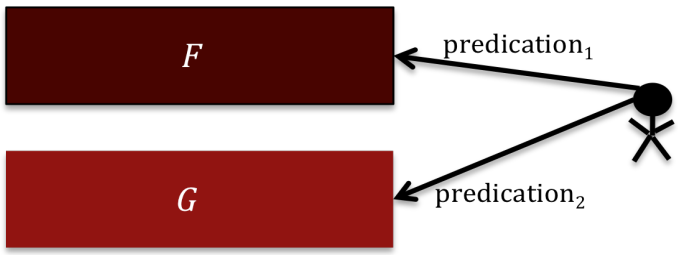

Figure 8

Intuitively, it is vague *which* property the perceiver depicted in Figure 1 predicates. After all, the perceiver predicates$_1$ F and predicates$_2$ G. Since it is vague whether 'z predicates y' expresses predication$_1$ or predication$_2$, there is no straightforward answer to the question: which property does the relevant perceiver predicate?

With these remarks in mind, consider the following 'toy model':

TOY MODEL

- There are exactly two predication-candidates: predication$_1$ and predication$_2$.
- There is exactly one phenomenal character: $\prod$.
- There are two properties $F \neq G$ such that:
  - a subject instantiates $\prod$ iff she predicates$_1$ F
  - a subject instantiates $\prod$ iff she predicates$_2$ G
  - no subject simultaneously predicates$_1$ and predicates$_2$ the same property

According to the toy model, if a subject bears any of the relations depicted in Figure 9, then she bears all of the relations depicted in Figure 9:

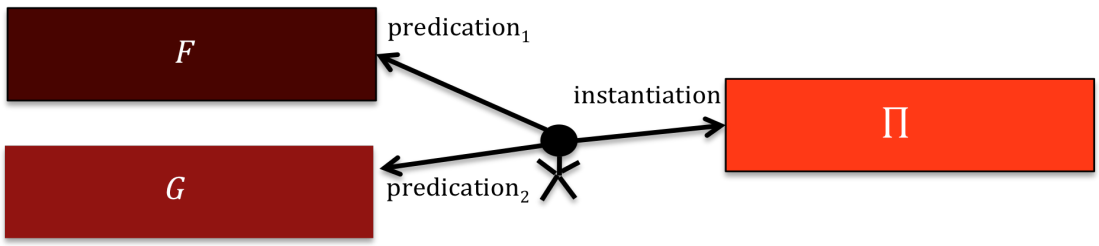

Figure 9

The toy model depicts a world in which representationalism *simpliciter* is true, yet Δ-representationalism is false. Begin by distinguishing the following claims:

5. $\exists y \forall x \Delta(x \text{ instantiates } \prod \leftrightarrow x \text{ predicates } y)$
6. $\Delta \exists y \forall x(x \text{ instantiates } \prod \leftrightarrow x \text{ predicates } y)$

(5) is an instance of Δ-representationalism. However, (6) is *not* an instance of Δ-representationalism: it is a strictly weaker claim, in which Δ takes wide scope with respect to every quantifier. If (5) is false in the toy model, then Δ-representationalism is also false; if (6) is true in the toy model, then representationalism *simpliciter* is also true.

Here are the standard supervaluationist truth-conditions for (5) and (6):

- (5) is true iff there is some property y such that for every predication-candidate R, a subject instantiates $\prod$ iff she bears R to y.
- (6) is true iff for every predication-candidate R, there is some property y such that a subject instantiates $\prod$ iff she bears R to y.

Granting these truth-conditions, the toy model guarantees that (5) is false and (6) is true—for precisely the same reason that ⌜$\exists n\Delta$(n is small & n+1 is not small)⌝ is false and ⌜$\Delta\exists n$(n is small & n+1 is not small)⌝ is true:

- In the toy model, (6) holds: for every predication-candidate R, there is indeed *some* property y such that a subject instantiates $\prod$ iff she bears R to y.
- Crucially, however, the relevant property is different with respect to each predication-candidate: *predication*$_1$ associates $\prod$ with F; *predication*$_2$ associates $\prod$ with G. In light of this fact, (5) fails in the toy model. For (5) is true only if both predication$_1$ and predication$_2$ associate $\prod$ with the *same* property.

At the beginning of this section, we asked: what must the world be like, in order for representationalism to turn out true and Δ-representationalism to turn out false? The answer to this question is now clear. If representationalism is true and Δ-representationalism is false, then the *toy model*—or something near enough—must accurately characterize the metaphysics of perception.[48] Of course, the real world may

[48] At this stage, I am assuming that *every* representationalist accepts (6), even if they reject (5):

5. $\exists y\forall x\Delta$(x instantiates $\prod \leftrightarrow x$ predicates y)
6. $\Delta\exists y\forall x$(x instantiates $\prod \leftrightarrow x$ predicates y)

differ in certain irrelevant respects from the toy model. The toy model assumes (for heuristic purposes) that there is only one phenomenal character and only two predication-candidates. These assumptions are obviously over-simplified. Nevertheless, if representationalism is true and Δ-representationalism is false, then the *structure* of perceptual predication must mirror the *structure* of the toy model: there must be multiple predication-candidates which associate the same phenomenal character with different predicable properties.

§2 Assessing the Toy Model

§2.1 The Problem of Free Recombination

In order to determine whether it is coherent to accept representationalism and deny Δ-representationalism, we must therefore ask: does the toy model provide a coherent account of the metaphysics of perception? (In answering this question, we shall ignore the toy model's merely-heuristic oversimplification.)

I confess to having encountered great difficulties when attempting to evaluate the cogency of the toy model. It is plainly a peculiar account of the metaphysics of perception, though it is challenging to articulate the source of its peculiarity. There are, however, two readily identifiable features of the toy model which create cause for concern. As a general rule, *distinct relations should be freely recombinable*. In other words:

FREE RECOMBINATION

In general, for any dyadic relations $R_1 \neq R_2$ and any objects $x \neq y \neq z$:

$$\Diamond[R_1(x, y) \ \& \ R_2(x, z)] \rightarrow \Diamond[R_1(x, y) \ \& \ {\sim}R_2(x, z)]$$

(Informally: if it is possible that x bears R_1 to y and x bears R_2 to z, then it is possible that x bears R_1 to y *without* bearing R_2 to z.) Thus, it is possible that Avril is behind Hal and in front of Martha, but also possible for Avril to be behind Hal without being in front of Martha.

The principle of free recombination is motivated by suspicion of *necessary connections between distinct relations*. If R_1 and R_2 are genuinely distinct relations, then there should not be modally brute connections between them: it should be

metaphysically possible to 'recombine' their relata in arbitrary configurations. R_1 and R_2 are like 'knobs on a stereo, in the sense that all combinations are possible' (Schaffer 2010: 352).[49]

I certainly do not intend to suggest that free recombination holds universally. But as a general rule, we should be very hesitant to postulate relations which violate it, without providing some special explanation which makes the brute necessary connections between the relevant relations appear less brute. Here is an example. Say that an person *P* is *numbered by n* iff *P* has *n* hands; say that *P* is *double-numbered by n* iff *P* has $n \times 2$ hands. Obviously, it is impossible for a person to bear the 'being numbered by' relation to the number 2 without bearing the 'being double-numbered by' relation to the number 4. Strictly speaking, this is a violation of free recombination. But it is a harmless violation. The relations 'being numbered by' and 'being double-numbered by' are trivially interdefinable, in a manner that renders the necessary connections between them wholly explicable.

The toy model generates violations of free recombination that appear far from innocent. According to the toy model,

1. $\forall x(x \text{ instantiates } \prod \leftrightarrow x \text{ predicates}_1 F)$
2. $\forall x(x \text{ instantiates } \prod \leftrightarrow x \text{ predicates}_2 G)$

Presumably, if the toy model is correct, then it is necessary:

3. $\Box\forall x(x \text{ instantiates } \prod \leftrightarrow x \text{ predicates}_1 F)$
4. $\Box\forall x(x \text{ instantiates } \prod \leftrightarrow x \text{ predicates}_2 G)$

By the transitivity of $\leftrightarrow$, (3) and (4) entail:

5. $\Box\forall x(x \text{ predicates}_1 F \leftrightarrow x \text{ predicates}_2 G)$

The toy model stipulates that $predication_1 \neq predication_2$. After all, the whole point of the model is to ensure there are (at least) two predication-candidates. The toy model

[49] Lewis (1986) and Armstrong (1980) contain classic discussions of similar recombination principles.

also stipulates that $F \neq G$. As such, (5) is a straightforward violation of free recombination. It is metaphysically impossible for a subject to bear one predication-candidate to a given property without also bearing a distinct predication-candidate to a distinct property. In effect, the toy model generates brute necessary connections between distinct predication-candidates.

Matters become even worse when we attempt to make the toy model more realistic. For simplicity, the toy model assumes that there is only one phenomenal character (namely $\prod$). In reality, however, there are countless phenomenal characters: $\prod_1 \ldots \prod_n$. If we extend the toy model accordingly, we will end up with an account that looks like this:

$\forall x(x \text{ instantiates } \prod_1 \leftrightarrow x \text{ predicates}_1\ F)$
$\forall x(x \text{ instantiates } \prod_1 \leftrightarrow x \text{ predicates}_2\ G)$

$\forall x(x \text{ instantiates } \prod_2 \leftrightarrow x \text{ predicates}_1\ H)$
$\forall x(x \text{ instantiates } \prod_2 \leftrightarrow x \text{ predicates}_2\ I)$

$\forall x(x \text{ instantiates } \prod_3 \leftrightarrow x \text{ predicates}_1\ J)$
$\forall x(x \text{ instantiates } \prod_3 \leftrightarrow x \text{ predicates}_2\ K)$
(*etc.*)

This account generates dizzying numbers of brute necessary connections between *predication*$_1$ and *predication*$_2$, in contravention of free recombination:

$\Box\forall x(x \text{ predicates}_1\ F \leftrightarrow x \text{ predicates}_2\ G)$
$\Box\forall x(x \text{ predicates}_1\ H \leftrightarrow x \text{ predicates}_2\ I)$
$\Box\forall x(x \text{ predicates}_1\ J \leftrightarrow x \text{ predicates}_2\ K)$
$\Box\forall x(x \text{ predicates}_1\ L \leftrightarrow x \text{ predicates}_2\ M)$
(*etc.*)

(When we take into account the fact that there are likely to be far more than two predication-candidates in the actual world, matters will become even worse.)

In the absence of any special explanation, which makes these brute connections appear less egregious, this is a strong reason for dissatisfaction with the toy model. Certainly, it is difficult to believe that *predication*$_1$ and *predication*$_2$ are interdefinable in a straightforward manner that explains the necessary connections between them.

§2.2 The Problem of Overdetermination

There is a further reason for dissatisfaction with the toy model. As it stands, representationalism is not an *explanatory* doctrine; it merely states that a biconditional relationship holds between phenomenal characters and predicated properties. But many philosophers also accept an explanatory version of representationalism (cf. Chalmers 2004):

EXPLANATORY REPRESENTATIONALISM

For every phenomenal character *x*, there is some property *y* such that a subject *z* instantiates *x* iff—<u>and wholly because</u>—*z* predicates *y*.

The relevant sense of 'because' is a matter for debate, but let us assume that it expresses *metaphysical explanation* or *grounding*.[50] It is common to distinguish between 'full explanations' of a given fact and 'partial explanations' of a given fact. As my usage of the term 'wholly because' indicates, I shall assume that explanatory representationalists hold that phenomenal facts are fully (not just partly) explained by representational facts.

Can proponents of the toy model accept explanatory representationalism? They can—but only if they decline to accept Δ-explanatory representationalism:

Δ-EXPLANATORY REPRESENTATIONALISM

For every phenomenal character *x*, there is some property *y* such that for any subject *z* it is determinate that: *z* instantiates *x* iff—<u>and wholly because</u>—*z* predicates *y*.

How should the toy model be enriched, to ensure that explanatory representationalism is true and Δ-explanatory representationalism is false? The only apparent possibility is this:

[50] See Fine (2010), Rosen (2010) and Schaffer (2010) for discussion of grounding.

THE EXPLANATORY TOY MODEL

- There are exactly two predication-candidates: predication_1 and predication_2.
- There is exactly one phenomenal character: $\prod$.
- There are two properties $F \neq G$ such that:
    - a subject instantiates $\prod$ iff—<u>and wholly because</u>—she predicates_1 F
    - a subject instantiates $\prod$ iff—<u>and wholly because</u>—she predicates_2 G
    - no subject simultaneously predicates_1 and predicates_2 the same property

For now-familiar reasons, the explanatory toy model guarantees that (1) is false and (2) is true:

1. $\exists y \forall x \Delta$($x$ instantiates $\prod$ iff—and wholly because—x predicates y)
2. $\Delta \exists y \forall x$($x$ instantiates $\prod$ iff—and wholly because—x predicates y)

As such, the explanatory toy model ensures the truth of explanatory representationalism while ensuring the falsehood of Δ-explanatory representationalism.

Unfortunately, the explanatory toy model also generates *explanatory overdetermination*. Whenever a subject instantiates a phenomenal character $\prod$, the fact that she does so is fully explained both by the fact that she bears a relation (predication_1) to F, and by the fact that she bears a distinct relation (predication_2) to G. Such explanatory overdetermination is arguably a Bad Thing Indeed.[51]

To be sure, some explanatory overdetermination may be harmless. It is often said that a disjunctive truth (e.g. P ∨ Q) is fully explained by each of its true disjuncts; when both its disjuncts are true, there is harmless explanatory overdetermination. (See Fine 2001: 22). However, the fact that a given perceiver instantiates a given phenomenal character seems not to be a disjunctive truth, in any sense.

[51] For opposition to explanatory overdetermination, see Kim (1993: 281), Lowe (2000: 572), and Schiffer (1987: 148).

Likewise, it is often said that an existential truth (e.g. $\exists xFx$) is fully explained by each of its true instances (e.g. Fa, Fb, etc.) When an existential truth has more than one true instance, there is harmless explanatory overdetermination. (See Fine 2010; Rosen 2010: 117). However, the fact that a perceiver instantiates a given phenomenal character seems not to be an existential truth, in any sense.

Likewise, it is often said that mental facts are fully explained by neutral facts, and also fully explained by microphysical facts. Such explanatory overdetermination is harmless, because neural facts themselves are explained by microphysical facts. In general, if P is fully explained by Q and P is fully explained by R, there is no problem if R itself is fully explained by Q (or vice versa). However, this principle does not help the proponent of the explanatory toy model. The fact that Avril instantiates $\prod$ is fully explained by the fact that Avril predicates$_1$ F and fully explained by the fact that Avril predicates$_2$ G. But surely the proponent of the toy model does not think that the fact that Avril predicates$_1$ F is fully explained by the fact that Avril predicates$_2$ G (or vice versa).

These are brief remarks: I do not intend to undertake an extensive detour into the explanation literature. For present purposes, it suffices to observe that proponents of the toy model who wish to adopt explanatory representationalism are forced to accept an ugly-looking form of explanatory overdetermination. In the absence of any special mitigating circumstances which suggest that the relevant explanatory overdetermination is harmless, this is a further reason for dissatisfaction with the toy model—especially to the extent that explanatory representationalism is an attractive extension of representationalism *simpliciter*.

…

Let us take stock. The fundamental question addressed throughout this section is: should representationalists accept Δ-representationalism? We began by considering the argument that it is permissible to insert Δ operators within representationalism wherever it is permissible to insert □ operators. This argument fails, for reasons addressed in **§1.1**. Subsequently, we asked: what must the world be like, in order for representationalism to turn out true and Δ-representationalism to turn out false? I have argued that the world must have structure depicted by the *toy model*, characterized

above. The question of whether representationalists must accept Δ-representationalism therefore reduces to the (more tractable) question of whether the toy model is cogent. We have explored two reasons for dissatisfaction with the toy model. First, it generates myriad violations of the principle of free recombination. To the extent that representationalists are keen to avoid such violations, they should accept Δ-representationalism and reject the toy model. Second, when the toy model is enriched to accommodate explanatory representationalism, it generates explanatory overdetermination. To the extent that representationalists are keen to avoid such overdetermination, and keen to accept explanatory representationalism, they should accept Δ-representationalism and reject the toy model.

The problem of vagueness developed in CHAPTER TWO is a genuine problem. It is wholly unclear whether representationalists are in a position to provide any adequate response. The only two serious options—adopting an ontology of vague properties, and denying Δ-representationalism—face difficulties. No argument in this chapter is advanced in a 'conclusive' spirit. Recent trends in the philosophy of vagueness and the philosophy of perception are at odds with one another; the tension between them merits further investigation.

This concludes our discussion of the problem of vagueness. The following chapter explores a new topic.

Dualism and Borderline Consciousness

CHAPTER FIVE

In previous chapters, we explored the relationship between vagueness and the representationalist theory of perceptual phenomenology. This chapter explores a new topic: the relationship between vagueness and the *dualist theory of consciousness*. I shall argue in favour of two claims:

A. It is plausible that consciousness admits borderline cases.
B. If 'strong dualism' is true, then consciousness does not admit borderline cases.

Collectively, (A) and (B) highlight a potentially unattractive feature of strong dualism.

We shall proceed as follows. §**1** argues in favour of premise (A). §**2** addresses the surprising number of contemporary philosophers who are tempted to deny (A). §**3** formulates strong dualism. §**4** and §**5** defend premise (B). Finally, the **CODA** discusses weaker versions of dualism which are compatible with borderline consciousness.

For rhetorical purposes, it is convenient to pretend that the central goal of this chapter is to develop an objection against strong dualism. However, my true interests lie in logical geography: who can, and who cannot, allow that consciousness admits borderline cases? This is a nontrivial question, worthy of exploration.

§1 Does Consciousness Admit Borderline Cases?

§1 Preliminaries

Consciousness admits *borderline cases* iff there is (or could be) a subject such that it is vague whether she is conscious. On my favoured analysis, a subject *A* is *conscious* iff:

$\exists y$(phenomenal(y) & A instantiates y)
There is some phenomenal character which *A* instantiates.

(See CHAPTER ONE for details). This section argues that it is plausible that consciousness admits borderline cases.

In light of our discussion of precise supervaluationism in preceding chapters, readers may find this assertion confusing. Doesn't precise supervaluationism entail that consciousness lacks borderline cases? After all, precise supervaluationism states that:

$\Box\sim\exists x\exists y\nabla(x \text{ instantiates } y)$

Necessarily, there is no object *x* and no property *y* such that it is vague whether *x* instantiates *y*.

Fortunately, the tension is apparent rather than real. To see why, let us distinguish two scenarios in which consciousness has borderline cases. In the first scenario, there is a subject *x* and a phenomenal character *y* such that it is vague whether *x* instantiates *y*. This scenario is plainly ruled out by precise supervaluationism: if it cannot be vague whether any subject instantiates any property, then trivially it cannot be vague whether any subject instantiates any property which is a phenomenal character.

But there is a second scenario in which consciousness has borderline cases, which is wholly compatible with precise supervaluationism. In the second scenario, there is a subject *A* and a property $\prod$ such that *A* determinately instantiates $\prod$, *but it is vague whether* $\prod$ *is a phenomenal character*. That is:

$\Delta(A \text{ instantiates } \prod) \ \& \ \nabla\text{phenomenal}(\prod)$

(Furthermore, assume that *A* does not instantiate any property which is determinately a phenomenal character in the relevant scenario.)

The compatibility of this scenario with precise supervaluationism is readily apparent. There is no object and property in the scenario such that it is vague whether the former instantiates the latter. Rather, the scenario countenances an object and a property such that the former determinately instantiates the latter, but it is vague whether the latter counts as a phenomenal character. Precise supervaluationism says

nothing to rule this out. Moreover, if the scenario in question obtains, then consciousness has borderline cases—for it is vague whether *A* is conscious. That is,

$\nabla\exists y$(phenomenal(y) & *A* instantiates y)
It is vague whether there is some property y such that y is phenomenal and *A* instantiates y.

Precise supervaluationism is therefore compatible with the view that consciousness has borderline cases. Here is another way of putting the same point. As we observed in CHAPTER TWO §**3**, precise supervaluationism entails that:

$\forall x$(phenomenal(x) → ~∇(Avril instantiates x))
For all phenomenal characters x: it is not vague whether Avril instantiates x.

This was a central premise in the 'the problem of vagueness' for representationalism, developed above. Precise supervaluationism cannot, however, establish the stronger claim that:

$\forall x$(~∇phenomenal(x) & ~∇(Avril instantiates x))
For all x: it is not vague whether x is a phenomenal character, and it is not vague whether Avril instantiates x.

That would be required in order to show that consciousness lacks borderline cases.

Although precise supervaluationism is consistent with borderline cases of consciousness, the foregoing discussion has revealed an important fact. *Consciousness can only have borderline cases in a supervaluationist setting if there is (or could be) a property x such that it is vague whether x is a phenomenal character*. We shall revisit this observation at several points in this chapter (cf. **§4**). Fortunately, there is no reason to rule out such properties *a priori*. In a physicalist setting, it is extremely plausible that there exist various physical properties such that it is vague whether they count as phenomenal characters. More on this shortly.

§1.2 A Sorites Series for Consciousness

So much by way of preliminaries. Is there any reason to believe that consciousness admits borderline cases? The answer appears to be positive.

Here is a standard way to show that a predicate F admits borderline cases. First, one constructs a finite series of times $<t_1 \ldots t_n>$, during which an object A undergoes gradual change along some salient dimension. The series is constructed to ensure that F determinately applies to A at t_1 and determinately does not apply to A at t_n. Subsequently, one raises the following question: where is the cut-off point for F in the relevant series? That is: which t_i is such that F applies to A at t_i, and does not apply to A at t_{i+1}? (Classical logic guarantees that there is such a t_i.) Now, there may be various obstacles to answering this question. But sometimes the question doesn't just seem difficult. It seems downright *embarrassing*. When the phenomenology of embarrassment rears its ugly head, there is good reason to suspect that F is vague: for certain t_v in the series, it is vague whether F applies to A at t_v.

Precisely such a series can be readily constructed for the case of consciousness. Consider an early-stage fetus, which is determinately not conscious. At some point several months later, the fetus is determinately conscious. Slice up the relevant period into n picosecond-long intervals $<t_1 \ldots t_n>$. A picosecond is 10^{-12} of a second, i.e. one trillionth of a second. A picosecond is to one second as one second is to 31,700 years. Moving at almost 300,000 metres per second, it takes around 3.3 picoseconds for light to travel one millimetre.

At which picosecond does the fetus become conscious? That is: which t_i is such that the fetus is unconscious at t_i, and conscious at t_{i+1}? (Since the fetus is not conscious at t_1 and conscious at t_n, classical logic guarantees that there is such a t_i.) This is an embarrassing question, analogous to questions like 'when does a person become bald?', 'when does a person become short?', 'when does a person become old?', and so forth. The natural conclusion is that there are certain picoseconds t_v such that it is vague whether the fetus is conscious at t_v—just as there are certain times at which it is vague whether a person is bald. Hence, it is possible that there exists a subject A such that it is vague whether A is conscious.[52]

[52] Here, I rely on the following reasoning: if it is possible that ⌜ϕ⌝ is true at a time t, then it is possible that ϕ. I doubt that anyone will object to this principle.

It may appear that future science will provide some criterion for determining whether a subject is conscious at a given picosecond. On reflection, however, this claim is very difficult to understand. Science is capable of revealing all sorts of facts about changes in the brain from picosecond to picosecond. (Though it seems unlikely that there is anything interesting to say. It takes around 3.3 picoseconds for light to travel one millimetre, moving at almost 300,000 metres/second. Sadly, impulses typically travel along neurons at a speed of anywhere from 1 to 120 metres per *second*.) Nevertheless, it is difficult to believe that science will help us identify the unique picosecond at which it is first appropriate to describe a fetus as 'phenomenally conscious'. Will future science will help us determine the unique picosecond at which it is first appropriate to describe a person as bald? Will future science help us determine the unique picosecond at which it is first appropriate to describe a person as old?

Let us assume, then, that there is some picosecond t_v such that it is vague whether the fetus is conscious t_v. If precise supervaluationism is true, this is *not* because there is a phenomenal character $\prod$ such that it is vague whether the fetus instantiates $\prod$ at t_v. Rather, there is a property $\prod$ such that the fetus determinately instantiates $\prod$ at t_v, but it is vague whether $\prod$ is a phenomenal character. In a broadly physicalist setting, it is easy to envisage—at least in crude terms—how this can be so. According to functionalism (e.g. Putnam 1967; Lewis 1966), every phenomenal character is identical to some functional property. Nevertheless, functionalists will surely agree that there are certain functional properties F such that it is vague whether F is a phenomenal character. (To obtain such a property, identify a functional property which is determinately a phenomenal character, and gradually subtract various components of its functional role. Eventually, you will reach a functional property such that it is vague whether it counts as a phenomenal character.) Perhaps the fetus under consideration instantiates such a functional property. Likewise, according to a strong version of type physicalism, every phenomenal character is identical to some neural state. Many properties of this type are determinately phenomenal characters, and many properties of this type are determinately not. (Some neural states are obviously irrelevant to phenomenal consciousness). However, it is plausible that there are certain neural states N_V such that it is vague whether N_V is a phenomenal character. Here is a simplified but suggestive example. To a first

approximation, the Crick-Koch hypothesis (1990) holds that a subject *A* is conscious iff neuronal oscillation of 40 MHz occurs within *A*. Consider the property of *undergoing neuronal oscillation of 39.87 MHz*. Perhaps it is vague whether the latter property is a phenomenal character. In the same vein, it has been proposed that a subject is in pain iff there is relevant activity in the pyramidal cells of layer 5 of the cortex involving reverberatory circuits.[53] Presumably, there is at least one type of 'borderline activity' *A* such that it is vague whether *displaying activity of type A in the pyramidal cells of layer 5 of the cortex involving reverberatory circuits* counts as a phenomenal character. (In a dualist setting, however, matters are more complicated. Much more on this later.)

§2 Diagnosing Dissent

In light of the sorites series constructed above, it is plausible that consciousness admits borderline cases. It is surprising, therefore, that a number of contemporary philosophers think differently. Colin McGinn writes:

> The concept of consciousness does not permit us to conceive of genuinely borderline cases of sentience, cases in which it is inherently indeterminate whether a creature is conscious: either a creature definitely is conscious or it is definitely not.[54]

For similar remarks, see Antony (2006), Campbell (1984), Chalmers (1996: 105), Copeland (1993: 257 n. 25), and Strawson (1994: 153). Papineau, who ultimately allows that consciousness has borderline cases, nevertheless regards his own view as counterintuitive (1993: 124).

We are faced with a puzzle. On one hand, the sorites series developed in §**1.2** provides *prima facie* motivation for accepting that consciousness admits borderline cases. Yet a number of distinguished recent philosophers deny—often in rather emphatic terms—that this can be so. What explains their dissent? What, in other

[53] Block and Stalnaker (1999).
[54] McGinn (1996: 14).

words, is the source of the apparently widespread intuition that consciousness cannot admit borderline cases?

It is sometimes said that if consciousness admits borderline cases, we must give up the 'common sense' view that consciousness is like an 'inner light' (Papineau 1993: 121; cf. Searle 1992: 3). I am not sure what it means to say that consciousness is like an inner light, and I strongly doubt that any such claim is part of common sense. In any case, the metaphor fails for all the wrong reasons. Everyone knows that there are borderline cases of (visible) light: it is borderline, for example, whether ultraviolet electromagnetic radiation with a wavelength slightly less than 400 nm is 'light' in the usual sense. Thus, if consciousness also has borderline cases, this does not conflict with the obscure idea that 'consciousness is like an inner light'.

Can we do any better? I submit that the central line of reasoning which underlies suspicion of borderline consciousness is simply this:

THE ARGUMENT FROM IMAGINATION

1. We cannot imagine a vaguely conscious subject.
2. Therefore, we have good reason to believe that it is impossible for a subject to be vaguely conscious.

(A subject is *vaguely conscious* iff it is vague whether she is conscious.) The argument from imagination has a very familiar structure: a given state of affairs is deemed unimaginable; on these grounds, the same state of affairs is deemed metaphysically impossible. Every opponent of borderline consciousness cited above makes some reference to the difficulties associated with 'imagining' or 'conceiving' a vaguely conscious subject. (See, for example, the McGinn passage quoted above.)

What should we make of the argument from imagination? Taken at face value, premise (1) appears simply to be false. Surely we *can* imagine a vaguely conscious subject: we can, for example, imagine a fetus developing over several months, which is vaguely conscious at some time during the relevant period.

However, there is a different way of understanding premise (1), which makes it rather more plausible. There are (at least) two ways of imagining a subject who instantiates a given property: one might imagine such a subject from an *external*

perspective, or one might imagine such a subject from an *internal* perspective. To illustrate, suppose that you receive the following instruction:

(I) Imagine a subject viewing a giant red canvas, which envelops her entire visual field.

There are two very different types of mental action that you might perform in response to this request:

- Roughly speaking, you may imagine *seeing* an ordinary subject who is standing in front of a red canvas with her eyes open. You thereby imagine a subject seeing redness from an 'external' or third-person perspective.
- Alternatively, you may imagine *being* a subject who is viewing a red canvas, *from the inside*. You thereby imagine a subject seeing redness from an 'internal' or first-person perspective.

The distinction between external and internal imagining is, of course, somewhat challenging to state clearly. Nevertheless, we have no trouble grasping the distinction[55].

With these remarks in mind, return to the argument from imagination. As we observed above, there is no obvious difficulty associated with imagining a vaguely conscious subject *from an external perspective*: visually depicting a fetus at some relevant stage of development appears to suffice. But there is a genuine difficulty associated with imagining a vaguely conscious subject *from an internal perspective*. Indeed, I am inclined to agree that this cannot be done. It is very difficult to imagine *being* a vaguely conscious subject. If premise (1) is understood to mean that we cannot imagine a vaguely conscious subject from an internal perspective, then its plausibility is considerable.

Nevertheless, the argument from imagination is a bad argument: from the fact that vaguely conscious subjects are unimaginable from an internal perspective, it does not follow that vaguely conscious subjects are *impossible*. There are far better explanations of the apparent unimaginability.

[55] For a recent discussion of internal imagination, see Martin (2002).

Let us begin with the following observation. It is easy to imagine a subject seeing redness from an internal perspective. But when we shift to alternative modalities, difficulties arise. For example, try to imagine a subject tasting water—or tasting Coca Cola—or smelling lavender—from an internal perspective. I find it extremely challenging to imagine such things, and I suspect that I am not alone: our capacity for olfactory internal imagination and gustatory internal imagination is limited at best. It certainly does not follow, however, that it is *impossible* to taste water, or to smell lavender: internal unimaginability does not entail impossibility. Furthermore, limitations relating to internal imagination are not confined to the olfactory and gustatory sphere. Even in the visual case, difficulties often arise. In the literature on 'non-conceptual content', it is often observed that ordinary subjects are unable to imagine seeing particular superdeterminate shades of colour from an internal perspective, even though it is plausible that we perceive such colours on a regular basis. Once again, internal unimaginability is a poor guide to impossibility.

There are other reasons to resist the argument from imagination. Consider the following analysis:

3. To imagine a vaguely conscious subject 'from an internal perspective' *just is* to imagine being introspectively aware of being vaguely conscious.

(3) is plausible. So is (4):

4. Necessarily: if a subject is introspectively aware of anything, then she is determinately conscious.

Introspective awareness is, after all, a high-level activity. Anyone engaging in such an activity is best categorized as determinately conscious. Vaguely conscious subjects do not engage in introspection or other sophisticated mental activity; their cognitive profile is very minimal indeed.

Together, (3) and (4) provide a compelling *independent* explanation of the fact that it is impossible to imagine a vaguely conscious subject from an internal perspective. By (3), to imagine a vaguely conscious subject 'from an internal perspective' *just is* to imagine being introspectively aware of being vaguely

conscious. But by (4), it is impossible to be introspectively aware of being vaguely conscious. In other words, to imagine a vaguely conscious subject from an internal perspective is to imagine something impossible. In general, we cannot imagine the impossible. It is no surprise, therefore, that we cannot imagine a vaguely conscious subject from an internal perspective.

This is a perfectly satisfying explanation of the fact that vaguely conscious subjects are unimaginable from an internal perspective. But the explanation does not rely on the claim that vaguely conscious subjects are impossible. Rather, it relies on the claim that it is impossible to be *introspectively aware of being vaguely conscious*. We have no reason, therefore, to draw the strong conclusions urged by McGinn and other opponents of borderline consciousness.

In summary: many contemporary philosophers are repelled by the suggestion that consciousness admits borderline cases. The source of their opposition is unclear. I have proposed a diagnosis of their intuitions, which identifies the argument from imagination as the primary motivation for denying that consciousness admits borderline cases. However, the argument from imagination is philosophically unmeritorious, and relies on an over-optimistic conception of the relationship between internal imagination and metaphysical possibility.

Our default assumption should be that consciousness admits borderline cases. This assumption is defeasible, like everything else in the philosophy of mind. But it has yet to be defeated.

§3 Strong Dualism and Fundamentality

§3.1 Introducing Strong Dualism

Is dualism compatible with the view that consciousness admits borderline cases?

There are many versions of dualism.[56] This chapter begins by developing a strong version of dualism. Weaker formulations are discussed below. In order to formulate dualism, we require the notion of a *physical* property. I shall not attempt to define 'physical' here. However, it is worth distinguishing between two conceptions of physicality. On the first conception, physicality is *closed under supervenience*: if

[56] It is common to distinguish 'substance dualism' from 'property dualism'. This chapter is concerned solely with the latter. The most influential recent discussion of dualism is Chalmers (1996).

$P_1 \ldots P_n$ are physical properties, then so is any property P^* which supervenes on $\{P_1 \ldots P_n\}$. So understood, both the property of being an electron and the property of being a table are physical properties. On the second conception, physicality is not closed under supervenience: only 'fundamental' properties count as physical. So understood, the property of being an electron is physical, but the property of being a table is not. I shall employ the first notion of physicality, which is closed under supervenience.

P^* supervenes on $\{P_1 \ldots P_n\}$ iff for any *accessible* worlds w, w^*: if w is a duplicate of w^* with respect to $\{P_1 \ldots P_n\}$, then w is a duplicate of w^* with respect to P^*. A world w^* is 'accessible' from a world w iff the same fundamental properties are instantiated in w and w^*.[57] (More on the notion of fundamentality shortly.) There are subtle questions pertaining to the correct analysis of the locution 'w is a duplicate of w^* with respect to $\{P_1 \ldots P_n\}$'; see Bennett and McLaughlin (2011) for discussion.

Strong dualism—on my formulation—comprises *four* related claims. The first two claims are these:

D1. $\forall x(\text{physical}(x) \rightarrow \Delta{\sim}\text{phenomenal}(x))$

D2. $\forall x(\text{phenomenal}(x) \rightarrow \Delta{\sim}\text{physical}(x))$

(D1) says: every physical property (and hence, every property which supervenes on the physical) is determinately not a phenomenal character. Likewise, (D2) says: every phenomenal character is determinately not a physical property (and hence, determinately does not supervene on the physical). Although (D1) and (D2) are strong claims, I suspect that they are endorsed by many card-carrying dualists. On my understanding, strong dualists also accept the following thesis:

D3. $\forall x\Delta(\text{phenomenal}(x) \rightarrow \text{fundamental}(x))$

According to **D3**, it is determinate that every phenomenal character is a *fundamental property*. The notion of fundamentality is extremely commonplace in post-Lewisian

[57] The importance of employing a restriction on 'accessible worlds' when defining supervenience is highlighted by Lewis (1983).

metaphysics. Lewis famously wrote:

> Because properties are so abundant, they are undiscriminating. Any two things share infinitely many properties, and fail to share infinitely many others. That is so whether the two things are perfect duplicates or utterly dissimilar. Thus properties do nothing to capture facts of resemblance. […] It would be otherwise if we had not only the countless throng of all properties, but also an elite minority of special properties.

We will refer to this 'elite minority of special properties' as *fundamental* properties. Lewis himself describes the relevant properties as 'perfectly natural'.

It is common to distinguish between a comparative notion of fundamentality (= *x* is more/less fundamental than *y*) and an absolute notion of fundamentality (= *x* is fundamental *simpliciter*). The notions are, of course, connected: *F* is fundamental *simpliciter* iff no property is more fundamental than *F*. This dissertation ignores the notion of relative fundamentality, concentrating instead on the absolute notion of fundamentality.

Fundamentality has been recruited to serve a variety of purposes in metasemantics, the study of duplication and intrinsicality, the metaphysics of natural laws, and much else besides.[58] There is no space to explore these matters here. One feature of fundamentality, however, is worth emphasizing. Fundamentality is connected to supervenience, in at least the following sense: *all properties supervene on the fundamental properties*. For this reason, it is plausible that **D2** entails **D3**. (It would be incoherent to hold that every fundamental property is physical, and also to claim that phenomenal characters do not supervene on the physical.)

The fourth (and final) commitment of strong dualism is this:

D4. $\forall x \Delta(\text{fundamental}(x) \rightarrow (\text{phenomenal}(x) \vee \text{physical}(x)))$

According to **D4**, it is determinate that every fundamental property is either phenomenal or physical. **D1**-**D4** guarantee that the set of fundamental properties is partitioned into *two* disjoint and collectively exhaustive classes: the class of physical

[58] See Sider (2011); Dorr and Hawthorne (ms).

properties, and the class of phenomenal characters. This puts the 'dual' in 'dualism'. (We can envisage alternative 'triadic' or 'quadratic' views which prefer to partition the set of fundamental properties into three or four disjoint and collectively exhaustive sets. Such views are discussed in the **CODA**.)

Crucially, I have not assumed that dualists accept the following claim:

(#) Every property is either physical or a phenomenal character.
$\forall x((\text{phenomenal}(x) \vee \text{physical}(x)))$

Indeed, dualists should deny **(#)**. To see why, let P_1 be a phenomenal character and let P_2 be a physical property. Consider the conjunctive property: *having P_1 and P_2*. This property is plainly not a phenomenal character. Nor is it a physical property, since (if dualism is true) it does not supervene on the physical. Therefore, at least one property is neither physical nor phenomenal. Similar remarks apply to the disjunctive property: *having P_1 or P_2*. **(#)** is therefore false. These counterexamples do not threaten **D4**, since properties like *having P_1 and P_2* and *having P_1 or P_2* are obviously not fundamental.

This concludes my (very brief) introduction to dualism. Although there is much left to say, the relatively limited scope of the present chapter allows us to dodge many tricky issues.

§3.2 Precise Fundamentality

Before we proceed further, I shall advance a claim about the relationship between fundamentality and vagueness. The claim will prove important when we argue (in **§4**) that strong dualism is incompatible with borderline consciousness.

PRECISE FUNDAMENTALITY

Every property is either determinately fundamental or determinately not fundamental.

Equivalently: there are no borderline or vague cases of fundamentality. Precise Fundamentality is extremely plausible. At any rate, the usual test for vagueness strongly suggests that Precise Fundamentality is true. It appears to be impossible to construct a 'sorites series' for fundamentality—a series of properties $<F_1 \ldots F_n>$ such

that:

- Each F_i differs ever-so-slightly from F_{i+1};
- F_1 is determinately fundamental;
- F_n is determinately not fundamental;
- There is no 'determinate cut-off point' in the series—i.e., no property F_j such that F_j is determinately fundamental and F_{j+1} is determinately not fundamental.

I simply cannot envisage any such sorites series. For this reason, there is a strong presumption that Precise Fundamentality is true. I readily concede that sorites-insusceptibility is not a *watertight* test for precision (Weatherson 2010)—but it is a very good indicator.

§4 Strong Dualism and Borderline Consciousness

Is strong dualism compatible with the view that consciousness admits borderline cases? The answer, I submit, is negative.

Suppose that strong dualism is true, and suppose—for *reductio*—that it is vague whether a subject *A* is conscious. In **§1.1**, we distinguished between two scenarios in which it is vague whether *A* is conscious:

SCENARIO ONE

There is a phenomenal character $\prod$ such that it is vague whether *A* instantiates $\prod$.

SCENARIO TWO

There is a property $\prod$ such that *A* determinately instantiates $\prod$, *but it is vague whether* $\prod$ *is a phenomenal character*.

As we observed above, only the second scenario is compatible with precise supervaluationism. (The first scenario countenances vague properties; the second scenario does not.) This dissertation operates under the assumption that precise supervaluationism is true. Accordingly, if it is vague whether a subject is conscious,

the second scenario obtains: there is some property $\prod$ such that it is vague whether $\prod$ is a phenomenal character. I shall argue in favour of the following claim:

PRECISION

If strong dualism is true, there is no property $\prod$ such that it is vague whether $\prod$ is a phenomenal character.

Evidently, Precision contradicts the twin assumptions that strong dualism is true, and that there is a subject *A* such that it is vague whether *A* is conscious. Conclusion: if strong dualism is true, consciousness does not admit borderline cases.

Why should we accept Precision? The crucial trick is to recall an important claim advanced in **§3.2**: that *fundamentality is precise*. The distinction between fundamental properties and non-fundamental properties is a precise one; fundamentality does not admit borderline cases. Strong dualists cannot countenance a property $\prod$ such that it is vague whether $\prod$ is a phenomenal character, for it would also be vague whether such a property is *fundamental*—contradicting the precision of fundamentality.

Begin by recalling the central tenets of strong dualism:

D1. $\forall x(\text{physical}(x) \rightarrow \Delta{\sim}\text{phenomenal}(x))$

D2. $\forall x(\text{phenomenal}(x) \rightarrow \Delta{\sim}\text{physical}(x))$

D3. $\forall x\Delta(\text{phenomenal}(x) \rightarrow \text{fundamental}(x))$

D4. $\forall x\Delta(\text{fundamental}(x) \rightarrow (\text{phenomenal}(x) \vee \text{physical}(x)))$

Using **D1**—**D4**, we shall argue that Phenomenal Precision holds. Let $\prod$ be an arbitrary property and suppose—for *reductio*—that it is vague whether $\prod$ is a phenomenal character. Since fundamentality is nonvague, one of the following conditions obtains:

A. $\prod$ is determinately not fundamental. That is, $\Delta{\sim}\text{fundamental}(\prod)$.

B. $\prod$ is determinately fundamental. That is, $\Delta\text{fundamental}(\prod)$.

A contradiction can be derived from both (A) and (B), assuming that strong dualism is

true.

With Respect To (A):

Suppose that $\prod$ is determinately not fundamental. By the dualist thesis **D3**,

i. $\Delta(\text{phenomenal}(\prod) \rightarrow \text{fundamental}(\prod))$

It is determinate that: if $\prod$ is phenomenal, then $\prod$ is fundamental.

In a weak modal logic, the following schema holds: $\Delta(\phi \rightarrow \psi) \rightarrow (\Delta{\sim}\psi \rightarrow \Delta{\sim}\phi)$.

Hence, (*i*) entails:

ii. $\Delta{\sim}\text{fundamental}(\prod) \rightarrow \Delta{\sim}\text{phenomenal}(\prod)$

If $\prod$ is determinately not fundamental, then $\prod$ is determinately not phenomenal.

By hypothesis, $\prod$ is determinately not fundamental. Therefore, (*ii*) entails:

iii. $\Delta{\sim}\text{phenomenal}(\prod)$

$\prod$ is determinately not phenomenal.

But (*iii*) contradicts our assumption that it is *vague* whether $\prod$ is phenomenal.

Conclusion: if $\prod$ is determinately not fundamental, a contradiction is true.

With Respect To (B):

Suppose, on the other hand, that $\prod$ is determinately fundamental. By the dualist thesis **D4**,

i. $\Delta(\text{fundamental}(\prod) \rightarrow (\text{phenomenal}(\prod) \vee \text{physical}(\prod)))$

It is determinate that: if $\prod$ is fundamental, then either $\prod$ is phenomenal or $\prod$ is physical.

By hypothesis, $\prod$ is determinately fundamental. By (K_Δ),

ii. $\Delta(\text{phenomenal}(\prod) \vee \text{physical}(\prod))$

It is determinate that: $\prod$ is phenomenal or physical.

We shall perform disjunction elimination on (*ii*). Suppose, in the first case, that $\prod$ is phenomenal. Then—by the dualist thesis **D2**—it follows that $\prod$ is determinately not physical. In a weak modal logic, the following schema holds:

$[\Delta(\phi \vee \psi) \,\&\, \Delta{\sim}\psi] \rightarrow \Delta\phi$

Thus, by (*ii*), it follows that $\prod$ is *determinately* phenomenal. Yet this contradicts our assumption that it is vague whether $\prod$ is phenomenal.

Suppose, on the other hand, that $\prod$ is physical. Then—by the dualist thesis **D1**—it follows that $\prod$ is *determinately not* phenomenal. Again, this contradicts our assumption that it is vague whether $\prod$ is phenomenal.

By disjunction elimination on (*ii*), we have an outright contradiction. Conclusion: if $\prod$ is determinately phenomenal, then a contradiction is true.

Summing Up

We began by assuming that it is vague whether $\prod$ is phenomenal. This assumption has been reduced to absurdity. Fundamentality is precise: either $\prod$ is determinately not fundamental (option (A)) or determinately fundamental (option (B)). But—assuming that strong dualism is true—both options lead to contradiction. I conclude that the claim labeled 'Precision' (above) is correct: if strong dualism is true, then there is no property $\prod$ such that it is vague whether $\prod$ is a phenomenal character. As we have seen, consciousness admits borderline cases only if there is some property $\prod$ such that it is vague whether $\prod$ is a phenomenal character. The moral is clear: if strong dualism is true, consciousness does not admit borderline cases.

§5 Vagueness and Fundamentality

How should strong dualists respond to the argument developed in **§4**? My argument relied heavily on the assumption that fundamentality is precise: for any property *F*, either *F* is determinately fundamental or *F* is determinately not fundamental. Might

strong dualists reject this claim?

In unpublished work, Cian Dorr and John Hawthorne discuss (without endorsing) several putative examples of vague fundamentality.[59] Consider the following list of properties:

LIST A

- *x is (temporally) before y*
- *x is necessary*
- *x is (spatially) to the left of y*
- *x is more fundamental than y*
- *x is a sphere*

Various philosophers have been (or might be) tempted to claim that one or more of the properties on list A is fundamental. Although I doubt whether any of the properties on list A are fundamental, let us suppose—for the sake of argument—that they all are. However, consider the following properties:

LIST B

- *x is (temporally) after y*
- *x is possible*
- *x is (spatially) to the right of y*
- *x is less fundamental than y*
- *x is a straight line*

Observe that the *n*th property on list A and the *n*th property on list B are interdefinable: *x* is temporally before *y* iff *y* is temporally after *x*; *x* is necessary iff the negation of *x* is possible; *x* is spatially to the left of *y* iff *y* is spatially to the right of *x*; and so forth. In spite of their indefinability, the properties on list A are plainly distinct from the properties on list B. The ordered pair <*e*, *e**> is in the extension of the property *after*; the ordered pair <*e*, *e**> is not in the extension of the property *before*;[60]

[59] See Dorr and Hawthorne (ms). Some of the examples discussed below are due to Dorr and Hawthorne; others are mine.
[60] Though the distinct ordered pair <*e**, *e*> is, of course, in the extension of the property *temporally before*.

therefore, by Leibniz's law, the relevant properties are nonidentical.

Anyone tempted to claim that the properties on list A are fundamental faces an embarrassing question: why not say instead that the properties on list B are fundamental? Conversely, anyone tempted to claim that the properties on list B are fundamental faces an embarrassing question: why not say instead that the properties on list A are fundamental? There is no reason, it seems, to regard the properties on one list as more metaphysically privileged than the properties on the other list, given their interdefinability. In light of this observation, Dorr and Hawthorne suggest that the following view has at least some appeal:

- It is determinate that either *before* is fundamental or *after* is fundamental (but not both). However, it is vague which disjunct obtains: it is vague whether *before* is fundamental, and it is vague whether *after* is fundamental.
- Likewise for every other pair of interdefinable properties from list A and list B.

This proposal exploits the fact that standard logics for Δ allow a disjunction to be determinate even when both its disjuncts are vague. The proposal captures the spirit of the idea that modal properties, temporal properties and so forth are fundamental, while dodging the embarrassing question of *exactly which* modal properties, temporal properties and so forth are fundamental. Evidently, if the proposal discussed by Dorr and Hawthorne is correct, then fundamentality is not precise.

Nevertheless, there is an alternative reaction. Why not simply hold that the properties on list A and the properties on list B are *all* determinately fundamental? Why should we accept the claim that either the properties on list A are fundamental or the properties on list B are fundamental, but not both? Dorr and Hawthorne suggest two potential motivations. The first motivation arises from the Lewisian view that fundamental properties are all *independent* from one another. In particular, no fundamental property supervenes on any other natural property. This 'independence requirement' implies that *before* and *after* cannot both be fundamental properties, since the former supervenes on the latter (and vice versa). The second motivation arises from Occamite considerations of parsimony. Other things being equal, if the set of properties deemed fundamental by a theory *T* is smaller than the set of properties

fundamental by a theory *T*+, then *T* is preferable to *T*+. By this criterion, the view suggested by Dorr and Hawthorne is preferable to the view that all the properties on list A and list B are fundamental.

These considerations are far from decisive. The motivation for the Lewisian independence requirement is unclear. In particular, it is not obvious what justifies the Lewisian requirement over a weaker requirement, to wit: any two fundamental properties are either independent or interdefinable. Yet this weaker requirement is consistent with the view that all the properties on list A and list B are fundamental. The motivation for the Occamite parsimony principle is also unclear. In particular, it is not obvious what justifies the Occamite parsimony principle over a weaker principle, to wit: other things being equal, if the set of non-interdefinable properties deemed fundamental by a theory *T* is smaller than the set of non-interdefinable properties deemed fundamental by a theory *T*+, then *T* is preferable to *T*+. Yet this weaker principle is consistent with the view that all the properties on list A and list B are fundamental.

In a sense, however, this is all a diversion. Let us assume—for the sake of argument—that the view discussed by Hawthorne and Dorr is correct. In certain special cases, it is vague whether a given property is fundamental. This does *not* affect our argument (developed in **§4**) that dualism is incompatible with borderline consciousness. We began the latter argument by assuming that there is some property $\prod$ such that it is vague whether $\prod$ is phenomenal. Subsequently, we claimed that it is not vague whether $\prod$ is fundamental, and derived a contradiction (using dualist premises). Is it appropriate to assume that it is not vague whether $\prod$ is fundamental? Perhaps there are certain properties *V* such that it is vague whether *V* is fundamental. However, this phenomenon only arises if there is some property *V** such that:

- *V* does not appear to be more metaphysically privileged than *V**;
- It is determinate that either *V* or *V** is fundamental; and
- *V* and *V** are interdefinable.

If the foregoing conditions obtain, say that *V** is a *mirror image* of *V*.

Prima facie, there is no reason to believe that $\prod$ has a mirror image. By hypothesis, $\prod$ is a property instantiated by a vaguely conscious human being. As

such, ∏ contrasts sharply with properties like *before*, *left*, *possibility*, and so forth, which possess mirror images. Since there is no reason to believe that ∏ has a mirror image, it is permissible to assume that it is not vague whether ∏ is fundamental.

…

Let us take stock. I have argued that if strong dualism is true, consciousness does not admit borderline cases. Consider, for example, the fetus discussed in §**1**. According to strong dualism, there is a unique picosecond t_d such that the fetus is determinately unconscious at t_d and determinately conscious at t_{d+1}. For reasons discussed above, I regard this as a deeply counterintuitive position.

Others, of course, will react differently. Indeed, the 'foes of borderline consciousness' identified in §**2** may even regard the argumentation in this chapter as an argument *for* strong dualism! Yet again, we face the eternal question: *modus ponens* or *modus tollens*?

That is a hard question—far above the pay grade of the present author.

Coda: Weak Dualism and Triadism

Not all forms of dualism conflict with the view that consciousness admits borderline cases. According to strong dualism,

D1. $\forall x(\text{physical}(x) \rightarrow \Delta{\sim}\text{phenomenal}(x))$

D2. $\forall x(\text{phenomenal}(x) \rightarrow \Delta{\sim}\text{physical}(x))$

Weak dualists, by contrast, assert neither **D1** nor **D2**. According to weak dualism, many physical properties are determinately not phenomenal characters, and many phenomenal characters are determinately not physical properties. But there are certain physical properties $P_1 \ldots P_n$ such that it is *vague* whether P_i is a phenomenal character, and/or certain phenomenal characters $P_1 \ldots P_n$ such that it is *vague* whether P_i is a physical property ($1 \leq i \leq n$).

Weak dualism is plainly consistent with the view that consciousness admits borderline cases; the argument in §**4** does not apply. On the other hand, weak dualism represents a significant concession. Say that dualism is *true of a phenomenal character P* iff *P* is nonphysical, and *true of a physical property P* iff *P* is nonphenomenal. Weak dualism concedes that there are certain phenomenal characters and/or certain physical properties of which dualism is not determinately true. It remains to be seen whether card-carrying dualists are happy to adopt this position.

There is another view in the vicinity—*triadism*—which also escapes the argument in §**4**. Triadists accept **D1**, **D2** and **D3**. But triadists do not assert **D4**:

D4. $\forall x \Delta(\text{fundamental}(x) \rightarrow (\text{phenomenal}(x) \vee \text{physical}(x)))$

Instead, triadists assert that there are three categories of fundamental property: the physical, the phenomenal, and the *exotic*. (Presumably, exotic properties are unknown to current science.) Crucially, triadists claim that there are various nonphysical exotic properties $E_1 \ldots E_n$ such that it is vague whether E_i is a phenomenal character ($1 \leq i \leq n$). On this view, a subject is borderline conscious iff she instantiates such an E_i. Triadism, like weak dualism, is consistent with borderline consciousness. Weak dualism is likely to repel many dualists on account of its conservatism. Triadism, by contrast, is likely to repel many dualists on account of its *extravagance*.

Chalmers (1996) discusses views which bear a close resemblance to weak dualism and triadism. Certainly, weak dualism and triadism deserve close attention. Unfortunately, not a single word remains to explore the matter further.

References

ADAMS, R. (1981) 'Actualism and Thisness'. *Synthese* 49: 3-41.

AKIBA, K. (2000) 'Vagueness as a Modality'. *The Philosophical Quarterly* 50: 359-370.

ANTONY, M. (2006) 'Vagueness and the Metaphysics of Consciousness'. *Philosophical Studies* 128: 515-538.

ARMSTRONG, D. (1980) *A Theory of Universals*. Cambridge University Press.

BARNES, E. AND WILLIAMS, R. (2009) 'Vague Parts and Vague Identity'. *Pacific Philosophical Quarterly* 90: 176-187.

BAYNE, T. (2010) *The Unity of Consciousness*. Oxford University Press.

BENNETT, K. AND MCLAUGHLIN, B. (2011) 'Supervenience'. In the *Stanford Encyclopedia of Philosophy*, edited by Zalta.

BLOCK, N. (1995) 'On a Confusion About a Function of Consciousness'. *Behavioural and Brain Sciences* 18: 227-47.

BLOCK, N. AND STALNAKER, R. (1999) 'Conceptual Analysis, Dualism, and the Explanatory Gap'. *The Philosophical Review* 108: 1-46.

BRAUN, D. AND SIDER, T. (2007) 'Vague, So Untrue'. *Nous* 41: 133-156.

BYRNE, A. (2001) 'Intentionalism Defended'. *The Philosophical Review* 110: 199-240.

BYRNE, A. (2009) 'Experience and Content'. *The Philosophical Quarterly* 59: 429-451.

CAMPBELL, K. (1984) *Body and Mind*. University of Notre Dame Press.

CHALMERS, D. (1996) *The Conscious Mind*. Oxford University Press.

CHALMERS, D. (2004) 'The Representational Character of Experience'. In *The Future of Philosophy*, edited by Leiter (Oxford University Press).

CRANE, T. (2008) 'Causation and Determinable Properties: on the Efficacy of Colour, Shape and Size'. In *Being Reduced*, edited by Kallestrup and Howhy (Oxford University Press).

CRICK, F. AND KOCH, C. (1990) 'Towards a Neurobiological Theory of Consciousness'. *Seminars in the Neurosciences* 2: 263-275.

COPELAND, J. (1993) *Artificial Intelligence*. Blackwell Press.

DIETZ, R. AND MORUZZI, S. (2010) *Cuts and Clouds* (editor). Oxford University Press.

DORR, C. (2010) 'Iterating Definiteness'. In Dietz and Moruzzi (2010).

DORR, C. AND HAWTHORNE, J. (ms) 'Natural Properties: a Guide for the Perplexed'.

DRETSKE, F. (1995) *Naturalizing the Mind*. Oxford University Press.

DUMMETT, M. (1975) 'Wang's Paradox'. *Synthese* 30: 301-24.

EHRING, D. (1996) 'Mental Causation, Determinables and Property Instances'. *Nous* 30: 461-480.

EHRING, D. (2011) *Tropes: Properties, Objects, and Mental Causation*. Oxford University Press.

EKLUND, M. (2011) 'Being Metaphysically Unsettled: Barnes and Williams on Metaphysical Indeterminacy and Vagueness'. In *Oxford Studies in Metaphysics volume 6*, edited by Bennett and Zimmerman (Oxford University Press).

EVANS, G. (1978) 'Can There be Vague Objects?' *Analysis* 38: 208.

FINE, K. (1975) 'Vagueness, Truth and Logic'. *Synthese* 30: 265-300.

FINE, K. (1985) *Reasoning With Arbitrary Objects*. Blackwell Press.

FINE (2010) 'Some Puzzles of Ground'. *Notre Dame Journal of Formal Logic* 51: 971-118.

FINE (2011) 'An Abstract Characterization of the Determinate/Determinable Distinction'. *Philosophical Perspectives* 25: 161-87.

FUNKHOUSER, E. (2006) 'The Determinate-Determinable Relation'. *Nous* 40: 548-569.

GRAFF, D. (2001) 'Phenomenal Continua and the Sorites'. *Mind* 110: 905-936.

HAMBLEN, C. (1973) 'Questions in Montague English'. *Foundations of Language* 10: 41-53.

HARMAN, G. (1990) 'The Intrinsic Quality of Experience'. *Philosophical Perspectives* 4: 31-52.

HAUG, M. (2010) 'Realization, Determination, and Mechanisms'. *Philosophical Studies* 150: 313-330.

HAWTHORNE, J. AND UZQUIANO, G. (2012) 'How Many Angels Can Dance on the Point of a Needle?' *Mind* 120: 53-81.

HECK, R. (1998) 'That There Might be Vague Objects (So Far as Concerns Logic)'. *The Monist* 81.

HILL, C. (2009) *Consciousness*. Cambridge University Press.

JOHNSON, W.E. (1921) *Logic: Part One*. Cambridge University Press.

JOHNSTON, M. (2004) 'The Obscure Object of Hallucination'. *Philosophical Studies* 120: 113-183.

KARTTUNEN, L. (1977) 'Syntax and Semantics of Questions'. *Linguistics and Philosophy* 1, pp. 3-44.

KIM, J. (1993) 'The Non-Reductivist's Troubles with Mental Causation'. In *Mental Causation*, edited by Heil and Mele (Oxford University Press).

LEWIS, D. (1966) 'An Argument for the Identity Theory'. *Journal of Philosophy* 63: 17-25.

LEWIS, D. (1983) 'New Work for a Theory of Universals'. *Australasian Journal of Philosophy* 61: 343–377.

LEWIS, D. (1986) *On the Plurality of Worlds*. Blackwell.

LEWIS, D. (1988) 'Vague Identity: Evans Misunderstood'. *Analysis* 48: 128-30.

LOWE, E. (2000) 'Causal Closure Principles and Emergentism'. *Philosophy* 75: 571-585.

LUDWIG, K. & RAY, G. (2002) 'Vagueness and the Sorites Paradox'. *Philosophical*

Perspectives 16: 419–461.

MARTIN, M. (2002) 'The Transparency of Experience'. *Mind and Language* 17: 376-425.

MCGEE, V. AND MCLAUGHLIN, B. (1995) 'Distinctions Without A Difference'. *Southern Journal of Philosophy* 33 (Supp.): 203–251.

MCGINN, C. (1996) *The Character of Mind* (2nd edition). Oxford University Press.

NAGEL, T. (1974) 'What is it Like to be a Bat?' *Philosophical Review* 83: 435-456.

OLSON, E. (2007) *What are We? A Study in Personal Ontology*. Oxford University Press.

PAPINEAU, D. (1993) *Philosophical Naturalism*. Blackwell.

PARTEE, B. (1986) 'Ambiguous Pseudoclefts with Unambiguous Be'. Reprinted in Partee's *Compositionality in Formal Semantics* (Blackwell).

PAUTZ, A. (2006) 'Sensory Awareness is not a Wide Physical Relation'. *Nous* 40: 205-240.

PAUTZ, A. (2007) 'Intentionalism and Perceptual Presence'. *Philosophical Perspectives* 21: 495-541.

PAUTZ, A. (2009) 'What are the Contents of Experiences?' *The Philosophical Quarterly* 59: 483-507.

PLANTINGA, A. (1974) *The Nature of Necessity*. Oxford University Press.

PUTNAM, H. (1967) 'The Nature of Mental States'. Reprinted in Putnam's *Mind, Language, and Reality* (Cambridge University Press).

ROSEN, G. (2010) 'Metaphysical Dependence: Grounding and Reduction'. In *Modality: Metaphysics, Logic and Epistemology*, edited by Hale and Hoffman (Oxford University Press).

ROSENTHAL, D. (2002) 'How Many Kinds of Consciousness?' *Consciousness and Cognition*.

RUSSELL, B. (1923) 'Vagueness'. *The Australian Journal of Philosophy and Psychology* 1: 84-92.

SCHAFFER, J. (2010) 'Monism: The Priority of the Whole'. *Philosophical Review* 119: 31-76.

SCHIFFER, S. (1987) *Remnants of Meaning*. MIT Press.

SCHIFFER, S. (2003) *The Things We Mean*. Oxford University Press.

SEARLE, J. (1959) 'Determinables and the Notion of Resemblance'. *Proceedings of the Aristotelian Society* 33: 141-158.

SEARLE, J. (1992) *The Rediscovery of theMmind*. MIT Press.

SIDER, T. (2011) *Writing the Book of the World*. Oxford University Press.

SIEGEL, S. (2010) 'Do Visual Experiences Have Content?' In *Perceiving the World*, edited by Nanay (Oxford University Press).

STANLEY, J. AND WILLIAMSON, T. (2000) 'Knowing How'. *Journal of Philosophy* 98: 411-44.

STRAWSON, P. (1994) *Mental Reality*. MIT Press.

TYE, M. (1990) 'Vague Objects'. *Mind* 99: 535-557.

TYE, M. (1995) *Ten Problems of Consciousness*. MIT Press.

TYE, M. (2000) *Consciousness, Color and Content*. MIT Press.

TYE, M. (2003) *Consciousness and Persons: Unity and Identity*. MIT Press.

TYE, M. (2009) 'Representationalist Theories of Consciousness'. In *The Oxford Handbook of Philosophy of Mind*, edited by McLaughlin and Beckermann (Oxford University Press).

VAN DALEN, D. (2008) *Logic and Structure*. Springer Publishing.

WEATHERSON, B. (2003) 'Many Many Problems'. *Philosophical Quarterly* 53: 481-501.

WEATHERSON, B. (2010) 'Vagueness as Indeterminacy'. In Dietz and Moruzzi (2010).

WILLIAMS, J.R.G. (ms) 'Classical Indeterminacy'.

WILLIAMS, J.R.G. (2008) 'Multiple Actualities and Ontically Vague Identity'. *Philosophical Quarterly* 58: 135-154.

WILLIAMSON, T. (1994) *Vagueness*. Routledge.

WILLIAMSON, T. (1999) 'On the Stucture of Higher-Order Vagueness'. *Mind* 108: 127-143.

WILLIAMSON, T. (2003) 'Vagueness in Reality'. In *The Oxford Handbook of Metaphysics*, edited by Loux and Zimmerman (Oxford University Press).

WILLIAMSON, T. (2004) 'Everything'. *Philosophical Perspectives* 17: 415-465.

WILSON, J. (2009) 'Determination, Realization, and Mental Causation'. *Philosophical Studies* 145: 149-169.

YABLO, S. (1996) 'How in the World?' *Philosophical Topics* 24: 255-286.