ZOMBIES AND THE ACTIVE IMAGINATION

By

CASEY WOODLING

A THESIS PRESENTED TO THE GRADUATE SCHOOL OF THE UNIVERSITY OF FLORIDA IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF MASTER OF ARTS

UNIVERSITY OF FLORIDA

2006
For my family.
ACKNOWLEDGMENTS

I owe a debt of gratitude to Professor Gene Witmer, which I hope to one day repay, for his insightful comments and suggestions and acute criticism, as well as his encouragement and patience in seeing the product to completion. I am also indebted Professor John Biro and Professor Murat Aydede for helpful comments and criticisms. Of course, none of this would have been possible without the support I received over the past year from my wife, Emily. Completing this project without her is simply inconceivable.
TABLE OF CONTENTS

ACKNOWLEDGMENTS........................................................................................................ iv

ABSTRACT ....................................................................................................................... vii

CHAPTER ............................................................................................................................

1 INTRODUCTION ........................................................................................................ 1

Zombies and Other Minds: A Brief Introduction......................................................... 1
Similarities Between Zombies and Other-mindless Creatures................................. 2
Differences Between Zombies and Other-mindless Creatures................................. 3
General Skepticism about the Conceivability-Possibility Link ................................. 5
The Conceivability of Zombies .................................................................................. 9
The First-Person/Third-Person Distinction and Eric Marcus’s Worry ..................... 10
A Final Introductory Note: An Important and Plausible Assumption ...................... 13

2 ZOMBIE-LIKE THOUGHT EXPERIMENTS ............................................................. 14

The Imitation Man and Awareness ........................................................................... 14
Ned Block’s Super-Blindsighter ................................................................................. 16
  Blindsight and Super-blindsight ............................................................................ 16
  Is the Super-blindsighter a Zombie? ...................................................................... 19
  Lessons from the Super-Blindsighter .................................................................... 20
Siewert’s Blindsight and Blursight .......................................................................... 21
Robert Kirk’s Zombie: Positive Evidence for Zombiehood .................................... 26
Intending, Supposition, and Stipulation in Imagination and Conceiving ............... 29
  Some Case Studies: Same Image, Different Representation ............................... 31
  More Case Studies: Imagining or Conceiving of Absence ................................. 33
Searle’s Silicon Brain and the First-Person Perspective in Imagination ................... 36
Zombie Dave ............................................................................................................. 37
Summary ...................................................................................................................... 40

3 A SKETCH OF A THEORY OF IMAGINATION/CONCEIVING ......................... 42

The Structure of Imagination .................................................................................... 43
  Situations ................................................................................................................. 43
  Target Situations and the Subject’s Description ..................................................... 45
Abstract of Thesis Presented to the Graduate School
of the University of Florida in Partial Fulfillment of the
Requirements for the Degree of Master of Arts

ZOMBIES AND THE ACTIVE IMAGINATION

By

Casey Woodling

August 2006

Chair:  D. Gene Witmer
Major Department: Philosophy

This essay is a partial examination of the nature of thought experiments that involve creatures whose behavior and mental lives diverge radically from the way behavior and mental lives are tied together in the actual world. I hope the examination sheds some light on these sorts of thought experiments, those somewhat mysterious episodes whereby we represent situations to ourselves to test their coherence or elicit intuitions about them. I think that many times in thought experiments we simply take it for granted that what we take ourselves to have succeeded in imagining is in fact that which we set out to imagine, without giving the exercise the proper attention. In what follows I try to spell out the process of imagining and conceiving in more detail, while drawing attention to some ways in which it can be misleading.

The target of this essay is a premise in David Chalmers's Zombie Argument: the premise that zombies are conceivable. Philosophical zombies are creatures identical to us in every respect, but they experience nothing. The conceivability of these creatures is taken to entail their possibility, which is used in a modal argument for an anti-physicalist conclusion.
In Chapter 1, I briefly discuss the status of this link between conceivability and possibility, as well as the similarities between zombies and creatures in the Other Minds Scenario. In addition to zombies and what I call creatures in the Other Minds Scenario, I also look at similar thought experiments (Chapter 2) in the hope that they can enrich our understanding of what the conceivability of zombies comes down to. Next (Chapter 3), I sketch an account of imagination. Last (Chapter 4) I try to draw some of the main threads of the essay together in the hope of convincing readers that we should approach the conceivability of zombies with more caution.
CHAPTER 1
INTRODUCTION

Anyone who thinks he can imagine that the moon is made of [green] cheese has a very sluggish imagination: the active imagination demands a pasture of antecedently necessary thousands and thousands of millions of cows, demands a way to preserve a piece of cheese in broiling heat, freezing cold, and vacuum for thousands of millions of years, demands some off-stage machina to protect a piece of cheese thousands of miles across from gravitational compressions into non-cheese . . . But any serious attempt to imagine the moon being made of green cheese must, like the unimaginable object itself, soon collapse under its own weight.

--Peter van Inwagen, “Review of The Coherence of Theism”

Zombies and Other Minds: A Brief Introduction

There was a time when philosophers were vexed by the problem of other minds, which arises when we start asking how we know that others are minded. Proceeding on the assumption that I know that I am minded, the problem is that the way I know this is different from the way I know that others are minded. Everyone has direct access to one’s own mental states, but only indirect access, via behavior, to the mental states of others. The threat of skepticism comes in when we face up to the fact that all we have to go on is this indirect evidence. Thus the question arises: if all we have to go on is observable behavior, can we be certain that others are minded? If one must be certain that p to know that p, a clear challenge to knowing that p in this case is the inability to rule out the possibility of a situation in which no one is minded but oneself. Call the hypothesis that correctly describes such a scenario “the other minds skeptical hypothesis.”

At present a problem is vexing philosophers that bears a family resemblance to the other minds skeptical hypothesis: the zombie hypothesis. They are not the same: in the zombie hypothesis not only are all the behavioral facts the same but also all the physical
facts. A philosophical zombie is a creature that is a mirror image of a creature in this world in terms of physical properties (and all those properties that supervene in virtue of sameness of physical properties) but lacks phenomenal properties. Let us call a property ‘phenomenal’ iff necessarily, if something instantiates it, there is something it is like to be the thing having it. Call a description of a scenario in which creatures instantiate all our physical properties but none of our phenomenal properties “the zombie hypothesis.” Many think that the situation that the zombie hypothesis describes is coherent, even if there are compelling arguments that physicalism is true. Likewise, many think that the other minds skeptical hypothesis is coherent, even if there are compelling arguments for the conclusion that we know that others have minds.

**Similarities Between Zombies and Other-mindless Creatures**

This coherence, it is argued, is enough to establish that zombies and creatures in the other minds scenario (hereafter “COMS”) are at least conceptually possible. While, the coherence of the respective situations is used to different ends, the creatures in both cases appear very similar. Zombies and COMS behave as if they are minded. They act just as actual conscious folks do. In fact the behavioral evidence we have to go on is strikingly similar; in fact, it is identical. It is granted that the creatures in question behave as we do when we know we are conscious. They discuss their experiences. They report their feelings. They do everything that we know we do when there is something it is like to be us. Despite

---

1 Fixing the physical facts fixes all other facts that supervene on them. Going up the hierarchy this includes chemical, biological, and so on.

2 Definition of “phenomenal property” is due to Professor Witmer, in conversation.

3 Type-B Materialists think that zombies are conceivable but not possible (Chalmers 1996). One strategy the type-B Materialist employs is to show that there are cases of conceivability that do not entail possibility (For an example see Aydede and Güzeldere 2005).
their repeated professions that they are phenomenally conscious, we seem to be able to doubt that zombies and COMS have minds.

**Differences Between Zombies and Other-mindless Creatures**

The first distinction I want to note is not about the creatures themselves but their respective arguments, so to speak. Let us note that the move from the conceivability of others lacking minds to skepticism about our knowledge of those minds is blocked in a rather straightforward way by an argument from analogy:

1. I know I have mental states.
2. I know that my mental states occur in a body of a certain type with certain types of attendant behavior.
3. Others have bodies and exhibit behavior analogous to my minded behavior.
4. Therefore, other people who satisfy this criterion in (3) are minded.

This inductive argument is enough for knowledge, provided that inductive reasoning is justified. But what it does not rule out is *the conceptual possibility* that our neighbors lack minds. Even so, we can live with that if we can rule out the skepticism. A similar move, though, cannot be made in the zombie case. In the zombie argument, the move is from conceptual possibility to metaphysical possibility. Even if we have good reasons to attribute phenomenal properties to alleged zombies based on an argument from analogy, as long as we can coherently doubt they have those properties, the argument (given in detail soon) seems to go through.

Another rather obvious difference between the two is the stakes. The other minds skeptical hypothesis, primarily, has an epistemological consequence. With the zombie hypothesis, there is more than epistemology at stake. The conceivability of zombies is used

---

4 While I address the putative distinction between conceiving and imagination later, in the Introduction I use the terms synonymously.
in a powerful argument against physicalism. Here is a crude formulation of that argument, call it the “Zombie Argument” (“P” stands for the conjunction of the physical truths and “Q” stands for a phenomenal truth):

1. \( P \land \neg Q \) is conceivable.
2. If \( P \land \neg Q \) is conceivable, then \( P \land \neg Q \) is possible.
3. If \( P \land \neg Q \) is possible, materialism is false.
4. Materialism is false (Chalmers 2002: 196)

If zombies are conceivable, then, the lesson is not merely epistemological: we learn something about the basic structure of the world. In light of this metaphysical conclusion, we should take care to be sure that what we believe we are conceiving of is in fact what we are conceiving of. After all, the stakes are no less than a fundamental truth about the world.

Another important distinction is that the evidence in the two cases is different in an important way. In the other minds case, we can imagine that the internal constitution of COMS is different. Because of this it is easier to conceive of other mindless creatures than zombies. We can tell a positive story that explains why—despite seemingly conscious behavior—those creatures lack minds. For example, as Eric Marcus (2004) notes, for all we know COMS may be radio-controlled marionettes or robots (since there is nothing stipulated about their internal constitution, as in the zombie case). This sort of story gives us some positive evidence that explains why (or helps to explain why) COMS lack consciousness. Even though their external behavior is identical to normally conscious creatures, we can imagine their insides being different in a way that explains the difference in experience. We can tell no such positive story about zombies, since they are our molecule-for-molecule

---

5 Within this view there is room for various positions, but for now suffice it to say that a physicalist is anyone who thinks everything is physical in the most fundamental sense: whether that means everything is, reduces to, or supervenes on physical stuff—we leave that open.
duplicates. In fact, when one begins reflecting on it, it is rather difficult to see how we can convince ourselves that creatures that are just like us in some many respects could lack consciousness. Answering this last sort of worry is a main goal of this essay. Just how is it that we can imagine zombies in the absence of the sort of positive story one can tell in the COMS case?

Another relevant distinction is that it is not clear that anyone need be conscious in the zombie world, whereas the intelligibility of the other minds problem requires a conscious individual. Many times the zombie hypothesis is described as a situation in which no one is conscious. Sometimes the zombie situation is described in a more minimal fashion with a least one conscious individual or as a situation phenomenally identical with the actual world save one phenomenal property. In Chapter 4, I raise a worry about a certain type of conceivability that seems to require that we imagine the zombie world as a world in which one person is conscious, which changes the way the scenario is typically described rather dramatically.

**General Skepticism about the Conceivability-Possibility Link**

Since the focus of this essay is (1) of the Zombie Argument, let us take a quick look at what it means to say that X is conceivable. First we may note that many philosophers do not consider zombies to be a serious empirical possibility: they are not worried that their neighbors lack mental states. What is at issue is the conceptual possibility of these creatures. The standard test for the conceptual possibility of X is whether we can coherently represent X to ourselves. If we can, then it is reasonable to think that X is possible, at least in some
sense. This link between conceivability and possibility is an old one. Descartes (1990) and Hume (2000) endorsed it. But is conceivability a reliable guide to metaphysical possibility?6

One negative response to this question is that the way we can coherently think has no bearing on how the world actually is: Thinking doesn’t make it so to sloganize the call of resistance. I think this natural response can be answered by the following line of thought. For starters, a negative point: If a situation is not conceptually coherent, then it is doubtful that it could be a metaphysical possibility. I do not think this claim requires much in the way of justification because I think many take it to be obvious.7 If we can rule out the possibility of anything, such elimination will start with logical and conceptual contradictions, so mere thought seems capable of giving us access to metaphysical truth. For instance, we rule out

---

6 It is difficult to define “metaphysical possibility” without packing too much in. One might say, for example, that metaphysical possibility is what is possible is relative to the real structure of the mind-independent world. Defining things this way, though, makes it hard to see how we have can guarantee that cognizers have any true access to modality. On the other hand, one might say that metaphysical possibility is conceptual possibility or logical possibility, but this looks to relative what should be a mind-independent modal category to the conceptual abilities of cognizers. Taking a stand on this is not necessary at this point. The question is merely supposed to elicit the realist intuition (which I do not necessarily endorse) that the way the world is could possibly be radically distinct from how we represent it and conceptualize it. I think Thomas Nagel offers a particularly clear representation of this. He writes,

the world is in a strong sense independent of our possible representations, and may well extend beyond them. This has implications both for what objectivity achieves when it is successful and for possible limits of what it can achieve. Its aim and sole rationale is to increase our grasp on reality, but this makes no sense unless the idea of reality is not merely the idea of what can be grasped by those methods. (Nagel 1986: 91)

7 Although it should be noted that some theists may disagree for the reason that they do not want to not limit God’s power by limiting the types of worlds he can create. Perhaps he could create worlds in which logical contradictions abound. Descartes seems to have had such a view.
the possibility of square circles on the grounds that we cannot even form a coherent
representation of them: the hypothesis is so unintelligible, it cannot even get off the ground.

Some authors (Tidman 1994) are skeptical about the traditional claim that
conceivability is a reliable guide to possibility. Others have expressed concern over the
mixing of different types of imagining (first-person and third-person) in thought
experiments used to generate “Cartesian intuitions” (Nagel 1974, Hill 1991). These sorts of
worries attempt to throw the traditional link into doubt. Tidman gives a statement of a
robust version of this sort of skepticism: “Accounts of a conceivability-possibility link must
answer the question: Why should we think there is any connection at all between what we
can or cannot conceive of and what is possible?” (307). This is an ultimately misplaced, but
important, question in need of an answer.

One response is that it simply misses the intuitive pull of the traditional accounts, on
which we ask ourselves about the possibility of a certain situation, and then using our ability
to conceive of it, attempt to pronounce to what degree it seems possible. Hume, for
example, thinks that our ability to imagine a golden mountain shows that it is possible, in
some sense, while our inability to imagine a mountain without a valley shows that the latter is
not. One good reason for accepting this traditional account is that it offers a non-trivial link
to modality. If not this, then what? What sort of positive access would we have to modal
knowledge? In the absence of the traditional link we would just have to say that, trivially,
actual facts are possible. This seems inadequate. It seems possible that I could have become
a copy writer and not a graduate student. Why? What provides evidence for this possibility?
Well, I can conceive of a situation in which I am busy writing copy for an advertising firm.

---

8 Type-B Materialists also think that zombies are conceivable but not possible. Many of the
Type-B approaches to the zombie argument rely on questioning the link between
conceivability and possibility.
Provided that there is no hidden logical or conceptual contradiction in this, then such a situation seems possible based on the fact that it is imaginable.

In this essay I do not intend to raise a general skepticism about this conceivable-possibility link. What I want to do is the opposite. I want to keep the link in good standing, because of the non-trivial access it provides to modality, by making some remarks that I hope will help convince some that we need to be more serious about our practice of conceiving. I think the link can be easily abused by philosophers, who offer up scenarios that are underdescribed and then say things like, “Surely such a situation is conceivable,” when in fact we do not have enough information about the situation in question to even properly understand what we are to judge.

Leaving the skeptical views, let us look at the dominant account of the link. Many take the view that situations that do not suffer from conceptual or logical contradictions are genuine metaphysical possibilities. This, I think, is a liberal view of metaphysical possibility. It allows that worlds with deviant laws of nature are possible. Some, however, deny such worlds. Necessitarians (Shoemaker 1998) reject the liberal view that the space of metaphysically possible worlds outstrips the space of natural or nomically possible worlds, largely for Kripkean reasons. Without entering this debate, we can still safely say that if our representation of a world contains a contradiction, then that world is metaphysically impossible. This is surely something on which the liberal and necessitarian views agree. But going further, I submit that it is not unreasonable to think that whatever is conceptually possible is metaphysically possible, necessitarianism notwithstanding.9

---

9 As Katalin Balog notes, Chalmers’s epistemology of modality is desirable because it can explain possibility and necessity in terms of our concepts. It does not need to appeal to ontologically questionable entities like possible worlds. She writes, “In this picture the truth
Often the phrases “conceptual possibility” and “logical possibility” are used interchangeably. For instance, the target of this essay, the zombie world, was first introduced by David Chalmers as being a logical possibility. This free interchange is somewhat frustrating, because it is not obvious that the sorts of possibility are, at bottom, the same. I propose we understand the relationship in the following way. First, note that logical possibility is fundamentally a relationship between variables and logical connectives. And conceptual possibility is, roughly, a relationship between concepts, propositions, and concept possessors, and the logical connectives “and,” “if,” “then,” “or,” “but” are not only words, but also express concepts we possess and use in reasoning. So, since variables and logical connectives express concepts, we can say that logical possibility is a species of conceptual possibility.

The Conceivability of Zombies

This brief discussion of conceptual possibility helps introduce and provide some background for the subject matter of this essay: the conceivability of zombies. David Chalmers “resurrected” zombies (1996) and argued that their logical possibility, along with some plausible premises, as we noted, shows that physicalism is false. His current view on the conceivability of zombies is more nuanced than his (1996). In his (2002), he introduces grades of conceivability that are linked with grades of possibility, but the basic idea has not changed.

What worries me about zombies is not that they stand in the way of an otherwise solid metaphysics of mind, but that it is not always clear what exactly we are to conceive of, or what philosophers take themselves to conceive of when they talk about zombies, and so it is makers of modal claims are not a realm of possible worlds, but rather facts about our concepts and ordinary empirical truths” (1999: 504).
not clear to me that we are conceiving of zombies in the right way. In this essay, I want to ensure that we have an accurate understanding of what it is we are trying to do when we conceive of zombies. I believe that conceivability intuitions themselves are under-examined and are treated in a much too cursory a manner.

In order to fix some terminology and try to isolate some of the issues involved, I look at thought experiments like the zombie hypothesis in Chapter 2. In Chapter 3 I take a look at imagination and conceivability in general. In Chapter 4 I attempt to pull some of the main threads together and apply them to a particular sort of conceivability that Chalmers holds is a way we can conceive of zombies and is a good guide to metaphysical possibility.

The First-Person/Third-Person Distinction and Eric Marcus’s Worry

Before moving on, we need to face up to what I think is a serious challenge to the conceivability of zombies. Eric Marcus articulates this worry in his (2004). In a nutshell, his worry is that to properly imagine zombies we must imagine that they lack of phenomenal properties, and to do this we must imagine them from a first-person point of view. But there is no first-person perspective to a zombie, so we are left to imagine them from a third-person point of view. But for Marcus this is not enough, since imagining the zombie world from a third-person point of view is exactly like imagining the actual world from a third-person point of view. We have no positive evidence that alleged zombies fail to instantiate phenomenal properties.¹⁰ This worry also highlights an important phenomenon that other authors mention in the context of conceivability: the inability to fill in crucial details (Tidman 1994, Chalmers 2002).

¹⁰ The contrary is the case. We have overwhelming evidence for thinking that they are conscious: they behave just as if they were.
The crucial detail for Marcus is subjective absence. We fail to positively imagine the lack of subjective features. According to Marcus, when we use the words ‘without consciousness’ to describe the zombie world, these words modify the act of imagination and not the object (the zombie world) (Marcus 2004: 477). We do not imagine a zombie world but misdescribe a conscious world by refraining from imagining zombies’ conscious mental states. Marcus thinks we need to imagine zombies’ lack of mental states in order to avoid misdescribing them as unconscious. This demands we imagine subjective absence, but this is impossible since any subjective view we take up in the zombie world will be one from which there is something it will be like. The problem with this, of course, is that there is nothing to imagine and imagining zombies’ lack of mental states requires us to imagine them from a first-person point of view which does not exist in the zombie world.

I think Marcus is right that the way the zombie world “looks” is the same as the way the actual world “looks”. I also think he is right that we are modifying the act of imagination. He, however, wants to say that this modification in the act stage, so to speak, cannot make for a modification in the object stage. I think this is wrong and in what follows I try to show that modifications in the act of imagination can carry over to the object of imagination (See Section on Supposition in Chapter 3).

For this argument to have any bite, there must be something to this distinction between first and third-person imagining. Here is Marcus’s definition of first-person imagining:

First-person imagining involves imagining what it is like to have various experiences. In such cases, we have an intuition of a world or a situation in which there is a certain configuration of objects and properties; but these objects and properties are understood from a subjective point of view. (Marcus 2004: 479)

I should also note that I think this distinction between act and object (or process and product) is crucial to discussing imagination.
This is not entirely helpful. First, it seems that, as defined, all acts of imagination are first-person acts, because any time we intuit a world or a situation our understanding of that situation is from our own point of view (whose else could it be from?). I suppose what Marcus is thinking, though, is that for an act of imagination to be a first-person one, the person who is imagining the situation must imagine that he is experiencing the situation.

Many thought experiments in the philosophy of mind trade on this distinction. Descartes’s argument, for example, for thought being essentially non-extended requires that we reflect on our own conscious mental states and ask whether or not we can conceive of them existing in the absence of extended things. Nagel’s argument against reductive physicalism in “What it is like to be a bat?” requires that one attempt to imagine what it is like to be a bat from the first-person based on the physical (third-person) facts of bat physiology. In any case, Marcus’s worry is important because it raises questions about whether what we have imagined is, in fact, a true zombie world or just a conscious world that we happen to not be paying enough attention to.

Before we move on, let us try to say something a bit more solid about first-person and third-person imagination, since the so-called distinction is not always clear what the difference is supposed to be. Many times the distinction is taken to be obvious. Chalmers, for instance, helps himself to the distinction without being explicit:

At most this shows that from the third-person point of view, my zombie twin and I are identical, so that you cannot be certain that I am conscious; but we know this all along. But it does nothing to imply that from the first-person point of view, I cannot know that I am conscious. From the first-person point of view, my zombie twin and I are very different: I have experiences, and he does not [emphasis mine]. (Chalmers 1996: 199)

What does it mean to say that imagining can be done from a third-person or first-person point of view? I take this up later, but for now we may draw an intuitive distinction: to say that a view is a first-person view means that that view involves a subject’s perspective
essentially in a way that a third-person view does not. For example, if I imagine a field of corn in the country, no subject is there essentially. Now, Berkeley was tempted to say that, no, there is a subject’s perspective there essentially, and that subject is the imaginer. But this is a confusion of the process of imagination with the product, or between imagining (act) and what one imagines (object). The act is the act of mentally representing the field and the object is the field itself. Nothing in the object involves a subject essentially, while the act always involves a subject essentially. To be from the first-person point of view, what one imagines must involve a subject essentially. And nothing in the object involves a perspective in a way that I have to imagine being a subject in the situation to properly imagine a field of corn. On the other hand, if I imagine what it is like to stand next to tall corn stalks, then that is a first-person act because the perspective is essential to the object. If I omit it, then the object is incomplete.

A Final Introductory Note: An Important and Plausible Assumption

In this paper, I maintain that phenomenal properties are not limited to sensations but are also instantiated by thoughts. While I do not defend this outright, I do discuss thought experiments in Chapter 2 that I think support the assumption. In any case, whether or not thoughts can be qualitative, we need to be clear of one thing. Whatever mental states zombies might be said to have—such as beliefs or other propositional attitudes—none of those states instantiate phenomenal properties. Zombies instantiate no phenomenal properties, despite being built like us, hooked up to the world like us, and behaving like us. There is nothing it is like to be a zombie.

---

12 I take the act-object distinction and the process-product distinctions to be synonymous.
CHAPTER 2
ZOMBIE-LIKE THOUGHT EXPERIMENTS

The Imitation Man and Awareness

Perhaps the first instance of the zombie idea in the contemporary literature is Keith
Campbell’s imitation man. In Body and Mind (1980), Campbell invites us to imagine the
following:

Suppose a being very like us except that instead of feeling pain when he burns his
finger or breaks his toe, he has no locatable sensations at all. He just spontaneously
gains a new belief, it just “pops into his head” that he has burned his finger or broken
his toe, as the case may be. Call this being an imitation man. His awareness of his own
body would be like our awareness that the car we are driving is getting a flat tire. Some
change in our body, of which we are not conscious, has a result that it just pops into
our heads that the tire is going flat. (Campbell 1980: 100)

And then later,

Think again of the imitation man, who duplicates all of a typical man’s acquisition,
processing, and retrieval of information, and all his activity, but for whom there are no
phenomenal properties. (Campbell 1980: 108)

The imitation man’s situation parallels the situation of the unlucky motorist who,
feeling that knowing bump, comes to realize his tire has gone flat. The imitation man does
not feel pain the way we do. In fact he does not feel pain at all but is aware of it in an
intellectual way. The fact that he is in pain just pops into his head.

While it is clear that the imitation man feels no pain, it is not clear he does not
instantiate any phenomenal properties, Campbell’s contrary claim notwithstanding. In fact, it
seems reasonable to conjecture that the imitation man is phenomenally conscious, because
he is aware of his beliefs.
Campbell says that the imitation man instantiates no phenomenal properties. But this seems to be based on the questionable assumption that phenomenal properties are limited to what philosophers call “raw feels.” This I think is too limited a view of what a phenomenal property is. It seems that being in a mental state of awareness of tissue damage is a state of awareness that surely must instantiate phenomenal properties.\(^1\)

Paying attention to the way Campbell describes the imitation man’s awareness by using the personal pronoun “his” makes it hard to see how a subject could be aware of a state in his body and yet fail to instantiate some phenomenal properties. Even if he does not instantiate the phenomenal property *being in pain*, he is aware that he instantiates that property. It seems that this awareness itself instantiates a phenomenal property.

Things might have been different, I suppose, had Campbell described this awareness in a more deflationary way. That is, had Campbell stipulated that the imitation man is *unconsciously aware* of his pain in a way that allows him report his pain without the thought “popping” into his head, then we could conceive of him as lacking all phenomenal properties. But this should worry us, as it depends on the paradoxical notion of unconscious awareness. How could he be aware of these pain judgments if they are unconscious mental states? This certainly strains coherence.

In the thought experiments we discuss, it is important to be on the lookout for the description of various creatures’ access to exogenous information. In many of the cases, the creature performs *as if* he had the exact same sort of access to the external world that we do, even though he does not. We will find that we can coherently imagine creatures in identical environments to us who fail to instantiate identical phenomenal properties, but it is not

---

\(^1\) This claim runs contrary to a common distinction between thoughts and sensations in the history of philosophy. Thoughts are typically thought to be non-phenomenal and sensations are thought to be phenomenal. I resist this distinction.
obvious that these creatures lack all phenomenal properties, the imitation man being a prime example.

The conclusion to draw here is that zombies cannot be aware of their judgments in the way that the imitation man was. If they are, I think it is obvious that they are phenomenally conscious. Of course, it is important to note that their experience would not have the same phenomenology as ours. The imitation man does not see a red flower as we see it. There is no visual phenomenal property he instantiates in virtue of seeing a red flower. Yet he is aware that there is a red flower in front of him and is phenomenally conscious in virtue his awareness.

**Ned Block’s Super-Blindsight**

**Blindsight and Super-blindsight**

In “On a Confusion on a Function of Consciousness” Ned Block tries to carve a distinction between the concept of access-consciousness and the concept of phenomenal consciousness (hereafter a-consciousness and p-consciousness). We should be somewhat comfortable with the meaning of “p-consciousness.” The meaning of “a-consciousness,” Block defines thus:

> A state is A-conscious if it is poised for direct control of thought and action. To add more detail, a representation is A-conscious if it is poised for free use in reasoning and for direct “rational” control of action and speech. (The “rational” is meant to rule out the kind of control that obtains in blindsight.) (Block, Flanagan and Güzeldere 1997: 382)

It is important to note that these two types of states are typically linked in the real world. If a state is a-conscious, then it is usually p-conscious and vice versa. And Block’s aim is not to dispute this. He wants to show that our concepts of these phenomena are distinct—a

---

2 By ‘visual phenomenal property’ I mean a property one instantiates when he has an experience as of a colored object.
distinction he thinks is neglected in the study of consciousness. To show this, Block asks us to imagine a super-blindsight. Super-blindsight is similar to regular blindsight in that in both an agent gets exogenous information from the external world but does not experience that information in the normal way. Blindsighters and super-blindsighters, like the imitation man, fail to share our phenomenology but have some access to the external world.

Super-blindsighters and blindsighters, however, are different in the following respect: supers are a-conscious in a way that regulars are not. Regular blindsight is not enough to show that p-consciousness and a-consciousness come apart. Block, in an entry for *A Companion to the Philosophy of Mind*, clearly states why:

Consider a blindsight patient who “guesses” that there is an “X” rather than an “O” in his blind field. The patient has no access-consciousness of the stimulus (because, until he hears his own guess, he cannot use the information freely in reasoning or in rational control of action), and it is plausible that he has no phenomenal consciousness of it either. (Guttenplan 1994: 215)

So, the important difference between regular blindsight and super-blindsight with regard to a-consciousness seems to be that the super-blindsight can prompt himself to “blindsee” and can hear his own guesses in a way that allows him to use the guesses as targets for reasoning. Let us take a look at Block’s fuller description of the super-blindsight, which sounds very much like the imitation man right down to the “pop”, to get a better idea:

The super-blindsight spontaneously says “Now I know that there is a horizontal line in my blind field even though I don’t actually see it.” Visual information from his blind field simply pops into his thoughts in the way that solutions to problems we’ve been worrying about pop into our thoughts, or in the way some people just know the time or which way is north without having any perceptual experience of it. The super-blindsight himself contrasts what it is like to know visually about an X in his blind field and an X in his sighted field. There is something it is like to experience the latter, but not the former, he says. It is the difference between *just knowing* and knowing via a visual experience. Taking his word for it, here is the point: the content that there is an X in his visual field is A-conscious but not P-conscious. The super-blindsight case is a very limited partial zombie. (Block, Flanagan and Güzeldere 1997: 385)
The information is a-conscious because the super-blindsighter can use it to reason about his environment in a way the regular blindsighter cannot. What makes the difference when it comes to the super having a-consciousness that the regular lacks is the amount and quality of the information we are supposed to imagine the super-blindsighter has about his environment and the fact that he can use his correct guesses as reasons. This, it seems, is enough to get the a-consciousness.

So, we are to imagine that the super-blindsighter is able to do normal things like play tennis and fix watches with the same or nearly the same ability as a normal sighted man. In imagining a man performing thus, do we not imagine a super-blindsighter? I think that most people would find this sufficient. We imagine a person playing tennis, fixing watches, and we convince ourselves that he is not p-conscious of the incoming information in a way that there is something it looks like to him for a tennis ball to appear in his visual field. We suppose that, in addition to the image of the man playing tennis and fixing watches, there are no phenomenal properties of a visual sort instantiated by his performance of these tasks. Is this sufficient for imagining a super-blindsighter? Additionally, we might also ask whether or not this is a sufficient for imagining a zombie?

I think the answer to both these questions is “no.” The first answer is “no” because we need to do more than just imagine a man playing tennis, fixing watches and then suppose that he lacks the phenomenal properties we associate with vision. What more is needed? We need to imagine that the accuracy of the super-blindsighter’s guesses is verified by independent means. Otherwise, we would not want to say that he reasons in a rational way. But unless we imagine that the super-blindsighter has a way to know that his “guesses” are accurate, unless we imagine the mechanism by which he comes to know his guesses are accurate in a reliable way, we have not imagined a-consciousness. Without this mechanism,
we clearly do not want to say that the super-blindsighter can use the content in a rational way, because the link between his mental states and the external world would be a complete mystery to him. Until we imagine that mechanism, we have not properly imagined him as being a-conscious.

Is the Super-blindsighter a Zombie?

It is tempting to think that the super-blindsighter is a zombie. Block does say that zombies are a-conscious but not p-conscious (Block, Flanagan and Güzeldere 1997: 387). I find this claim implausible, even though it was responsible for my past belief that super-blindsighters were zombies. But it is important to see that the super-blindsighter is not a zombie, at least not a complete zombie. The super-blindsighter will instantiate phenomenal properties for the same reason the imitation man will. The content that is a-conscious will also be p-conscious in some sense. Let me explain. Block himself grants that the super-blindsighter’s thought of x is both a-conscious and p-conscious:

Of course, the super-blindsighter has a [thought] that there is an x in his blind field that is [both] A-conscious and P-conscious. but I am not talking about the thought. Rather, I am talking about the state of his perceptual system that gives rise to the thought. It is this state that is A-conscious without being P-conscious. If you are tempted to deny the existence of these states of the perceptual system, you should think back to the total zombie just mentioned. Putting aside the issue of the possibility of this zombie, note that on a computational notion of cognition, the zombie has [all] the same A-conscious contents that you have (if he is your computational duplicate). A-consciousness is an informational notion. The states of the super-blindsighter’s perceptual system are A-conscious for the same reason as the zombie’s. (Block, Flanagan and Güzeldere 1997: 385-6)

Block grants that thoughts can have p-conscious aspects. But is it not the thought of x that is the content poised for rational action? It is not, according to Block, the thought of x

3 In fact, I thought this at one time myself. It is a way to explain how it is that zombies can act as we do without sharing our phenomenology. The super-blindsighter will be a sort of partial zombie if we stipulate that he fails to instantiates all the phenomenal properties of a fellow who fixes watches and plays tennis in the actual world.
that shows that the two states of consciousness come apart. It is, rather, the state of the perceptual system that is poised for rational action that gets inside the head without instantiating any phenomenal properties (until the super is aware of it, of course). It carries information that is a-conscious but not p-conscious (until the super is aware of course). It is perhaps interesting to note that the super’s a-consciousness and p-consciousness come apart only relative to a certain cognitive system. It seems that this allows us to still retain the notion of awareness but separate the two senses of consciousness relative to a system. It is something else entirely, as in the zombie case, to say that there is a cognitive system that mirrors ours in terms of a-consciousness but is p-conscious of none of that information.

Lessons from the Super-Blindsighter

To sum up: Block thinks that the sense in which the two concepts come apart is only relative to a particular phenomenal property, say, the experience of seeing the red line in one’s blind field (p-consciousness) and being able to use that information to reason about the world (a-consciousness). He thinks that the easiest way to show that the concepts come apart is in the super-blindsight case. While he suggests that zombies are a-conscious but not p-conscious, he is hesitant to say much about the gross separation of a-consciousness and p-consciousness in the zombie world. Further, one can consistently think that the concepts of a-consciousness and p-consciousness are distinct, and deny they can come apart in toto. If this is true of thinking of zombies, then we have a way to explain why zombies seem prima facie conceivable but are not genuinely conceivable because of certain traces of p-consciousness. For example, if zombies were lacking in visual phenomenal properties that
would not be enough to show that they lacked all phenomenal properties.\footnote{It would be enough to show, if we stipulated that all the physical facts were the same, that the phenomenal facts were not necessitated by the physical facts. But often times the zombie world is described as a world in which no-one is p-conscious.} We also must coherently represent them as \textit{never being} p-conscious of their a-conscious states. This is surely a difficult task, one which Block refrains from developing a case for. He merely states that zombies are “A-conscious, but not P-conscious” (Block, Flanagan and Güzeldere 1997: 387).

**Siewert’s Blindsight and Blursight**

In \textit{The Significance of Consciousness} (1998) Charles Siewert attempts to make the concept of phenomenal consciousness clear via a series of thought experiments involving blindsight. The basic idea is that while the blindsighters he mentions (who are in various degrees of visual disrepair) behave as if they instantiated normal phenomenological properties, they cannot properly be called normally phenomenology conscious because they do not experience the world as we do. If one claims that the blindsighters are conscious in all the relevant ways, then he is neglecting an essential aspect of consciousness, what Siewert calls “phenomenal character.”

While the exercise is instructive and offers a more textured understanding of phenomenal consciousness than the standard fare, what concerns me is not the lesson as much as a particular thought experiment, and Siewert’s instruction that we are to imagine the beings therein from a first-person point of view. In this section I hope to give a better idea of what first-person imagining comes to and to address some skepticism regarding its use in thought experiments.

Consider one of the examples Siewert gives (1998: Chapter 4, Section 3). The example involves blindsight and amblyopic vision. Amblyopic vision, unlike blindsight, involves what Siewert calls “consciously seeing,” seeing in a way that involves phenomenal properties of a
visual sort. Amblyopic sighters have a visual presentation of the world such that there is something it is like to have that presentation, but that presentation is severely downgraded in an informational sense. An amblyopic sighter experiences blurry colors and no definitive shapes, leading Siewert to call this “blursight.”

Siewert asks us to imagine being an amblyopic named Connie who sees blurry patches in her left visual field and has normal vision in her right visual field. She cannot extract fine-grained information from her left visual field because of the blur sight. For example, she has information, with matching phenomenology, of the sort that tells her there is such and such shade of gray in front of her but it course-grained information. It is, like the minimal information the blindsighter (not the super of course) gets, no more definite than as of a shape like an “X” or an “O.” Next Siewert asks us to imagine being Belinda, a typical self-prompting blindsighter, who gets information from her environment that is identical in informational terms with the sort Connie gets. Intuitively, it seems that I can have a good idea of what it is like to be each person, in the sense that I can imagine being in the same mental states as those described by Siewert.

I can imagine what it is like to have the amblyopic sight. My visual experience is extremely fuzzy. At best, I can discriminate patches in my environment. I also seem to be able to imagine what it is like to be Belinda. But this seems to be important in a different way. In the case of Connie, I have a rough model to work from. Her experience is like mine but much more blurry. However, what might it be like for Belinda to experience the world? I suppose it is very much like imagining a super-blindsighter except for not getting a mental read-out that is super-rich in informational terms; rather, it is information-poor. Instead of desk: 48 x 60 inch top, 3 feet tall, et cetera, I get brown blur. What is similar is the darkness inside, which is only punctuated by information that just pops into my head.
So far that seems sufficient to achieve our goal. However, Witmer (2001) presses this sort of claim about first-person imagination. He argues that the first-person requirement here is too restrictive, because it seems to rule out imagining extra-mental conditions:

If the target of the imaginative exercise is limited to the imagining of having certain experiences, it hardly seems sensible to include as part of the requirements on imagining such a thing that you succeed in imagining that various extra-mental conditions obtain. (Section 3.5)

The point is that in many cases we need to imagine third-person features in order to achieve the goal, and to do that we must imagine from a third-person point of view. Witmer uses the case of imagining being Descartes sitting in front of his fire and imagining being Descartes dreaming that he is sitting in front of his fire. If we imagine both of these from the first-person perspective there is no distinction. They are experientially, or first-personally if you will, identical. What we need is a way to imagine those situations from a third-person perspective as well, in order to ensure that we are achieving our goal. But this gives rise to a further objection.

This objection is made by Christopher Hill (1991). He argues that many of our “Cartesian Intuitions” are fueled by splicing together products of two very different types of imagination: the first-person and the third-person, or what he calls, respectively, “sympathetic” and “perceptual” following Nagel (Nagel 1974). When we imagine, for example, C-fibers firing but no pain, we simply imagine a soggy grey brain with neurons lit up in the right configuration and then from the first-person perspective imagine being in a non-pain state. Then we splice these two images together in our mind and take ourselves to have achieved the goal of imagining a case of C-fibers firing but no pain. Hill thinks that we have good reasons to doubt that the mechanism underlying this phenomenon of splicing is a reliable guide to modal knowledge. He thinks that these cases are of the water-H\(_2\)O sort, where we misdescribe an impossibility as possible. Hill writes,
Consider intuitions to the effect that water can exist without being accompanied by H$_2$O. It is plausible to attribute such intuitions to a mechanism that operates by splicing together images of the presence of water and images of the absence of H$_2$O. (Hill 1991: 70)

In short, when we perform this splicing we are duped. Let us consider how this sort of response can be made against the Belinda and Connie case.

To properly address Witmer’s worry we need to imagine third person features. So we imagine Connie standing in front of a red fire-hydrant, pointing at it, and describing it as “Darker blob on top of lighter fuzzy stuff in front of even lighter fuzzy stuff” from a third-person point of view. Next we imagine Connie from a first-person point of view: we just imagine fuzzy figures, in various shades of grey, that in some very remote way resemble a fire-hydrant. But note: these seem to be two distinct products of imagination. How can we be sure that they are properly attributed to the same total product? Remember your Hill: this involves the unreliable mechanism of splicing. Further, when we repeat the exercise for Belinda, we have the same problem. And further when we compare the four acts we are unsure whether or not we have got a case of two creatures who are the ones we set out to imagine, or four separate creatures, the two we imagine from the third-person view, and the two we imagine from the first-person point of view.

I think this is overly restrictive. But it does point out an important issue. We might introduce a notion here—call it “tracking”—to help make sense of this phenomenon where we imagine the same situation from different perspectives. How does one ensure that proper tracking is going on? It seems that I do something like suppose that Connie is the same person whom I imagine from the two perspectives. I do not think we need to cast this sort of practice in a negative light. Why? Because, intuitively, it is the way we ensure that what we take ourselves to imagine is in fact the thing we are imagining in many other cases. So to cast it into doubt is to cast into doubt a large set of intuitions, which seems overly restrictive. For example, if Hill is correct, we could rule out a wide variety of imaginative episodes. For example, episodes that represent skeptical scenarios are ruled out because they depend on imagining our experiences being the same (first-person) and the external world being radically different (third-person), and then splicing these together. It certainly seems that these sorts of situations are conceivable.
The blursighter also shows us something interesting about the aforementioned distinction between a-consciousness and p-consciousness. If we can have a clear conception of what it is like to be the blursighter, then we have a case of having p-consciousness without our typical a-conscious states. There is something it is like to see things in the left visual field for the blursighter, even though the a-consciousness is well below what typically accompanies p-consciousness. Block also gives a case of p-consciousness without a-consciousness:

Suppose that you are engaged in intense conversation when suddenly at noon you realized that right outside your window, there is—and has been for some time—a pneumatic drill digging up the street. You were aware of the noise all along, one might say, but only at noon are you consciously aware of it. That is, you were P-conscious of the noise all along, but at noon you are both P-conscious and A-conscious of it. (Block, Flanagan and Güzeldere 1997: 386-387)

This case suggests that we might be in a phenomenal state that is blocked from our a-conscious system. In any case, something like this might happen in cases where people are in pain but are in the sort of situation where their a-conscious states are so focused on other things (life and death emergencies) that while they instantiate phenomenal properties that are not targeted or integrated into the a-conscious system. There is something it is like to have them, even though the subject is not aware of those properties.

Connie and Belinda also show that the same p-conscious state need not accompany the same a-conscious state. This is a familiar line of thought. It is essentially what grounds the inverted spectrum objection to functionalism. Although, Belinda and Connie are not in the same functional state, they are in the same a-conscious state. And if someone held that p-consciousness is reducible to a-consciousness, it looks like Connie and Belinda would be a compelling counter-example to that claim.
Robert Kirk’s Zombie: Positive Evidence for Zombiehood

In “Sentience and Behavior” Robert Kirk describes a man who becomes a zombie over a period of time. Here is the first stage:

One day Dan (who had been perfectly normal until that time) accidentally cut his hand and started to behave very strangely. True, his behavior included features indistinguishable from the typical behavior of a normal person suffering pain of the kind and intensity to be expected in the circumstances. He winced, exclaimed “Ouch!” nursed his hand, answered “Yes” to questions about whether it hurt, and so on. But in addition to, and superimposed upon these normal bits of behavior was a second kind of behavior wildly different from the typical behavior of someone in pain. He expressed astonishment; and he uttered sentences and behaved in ways which, in more normal contexts, would have been taken to be protestations to the effect that he felt no pain whatever, that it was as if he were “totally anesthetized.” He appeared to be astonished by two things: first, by the fact (as he seemed to regard it) that he really felt no pain at all in spite of a fairly serious injury; and second, by the fact that he was nevertheless wincing, groaning, and uttering such sentences as “It hurts like hell.” (Kirk 1974: 44)

So, Dan says he feels no pain even as he exhibits pain-behavior. He utters sentences like “That hurts like hell,” but he also utters sentences like “I do not know why I just said ‘that that hurts like hell,’ since I did not actually feel anything.” Kirk asks us whether this type of behavior is intelligible.

While he admits that is extremely puzzling, he does think it is intelligible, provided we assume that Dan feels no pain. It seems difficult to come up with an explanation of why Dan acts the way he does. But that is not Kirk’s target.

Perhaps, one way to make sense of it is to think of this occurring in a world where parallelism is true: mental states only appear to cause behavior. In fact they are simply designed this way by God. And suppose further that God has decided to be kind to Dan and not allow him to feel pain. God has designed things in such a way that those phenomenal properties from Dan’s mental states that are parts of mental substance when Dan is in pain states. Of course, his body is stimulated in such a way that it reacts in a normal fashion, it is just that—per God’s design—he is not to feel pain at certain times. This seems like a coherent story that we might tell to make Dan’s behavior a bit more intelligible.
Dan experiences further strangeness. He loses not just his ability to feel pain, but starts to lose his sensory modalities one by one. The curious thing is that he still acts as if he saw, heard, tasted, and smelled the world, but he also notes that he is not experiencing what he normally does, despite the fact that he is behaving as if he did—the first-person reports to the contrary notwithstanding.

Dan wakes one morning and, with much effort, manages to announce that he has gone blind. He struggles to tell the world that he is losing his conscious connection to it. His consciousness is fading out, while his behavior is as normal as can be. Then, after some time, Dan stops protesting. He just acts as he did in the beginning, before all the strange goings-on. Is he a zombie or not? Well, Kirk seems to think we have reasonable indirect evidence for believing that Dan is a zombie, even if we cannot directly conceive of him being one in the sense that we can positively imagine him not being conscious (cf. Marcus’s worry). What we can imagine is all the build up, not the missing consciousness itself.

Kirk sets things up in such a way that Dan actually comes back to his body, so to speak. After some time of normal behavior by the body that used to be Dan (Kirk describes this creature as “D”), that creature starts to assert that he can see again. Just as he lost his senses, Dan slowly regains them. This gradual return to normality suggests that D was a Zombie and not a person at all. But why should we rule out the possibility that D is conscious? D behaves as if he were: in fact, he reports none of the bizarre mental states that Dan does. Surely we can be justified in believing that Dan is no longer conscious, but what sense does it make to say that D, the new “person,” is a Zombie? To this Kirk writes,

---

7 I suppose that Kirk could have made the point without having Dan come back, so to speak. Yet, Dan regaining his consciousness allows us to better “frame” the zombihood between the strange goings on.
I think I have said enough to make it clear that the hypothesis that Dan had been superceded by another person would be less plausible than the hypothesis that he had turned into, or been superceded by, a Zombie; and faced with the sorts of behavior that I have ascribed to Dan, we should be justified in supposing that, for a time we had been dealing with a Zombie. (Kirk 1974: 50)

The main reason that Kirk thinks that D is unconscious is that if he were conscious, then he would give some indication of the puzzlement in which he finds himself, given that he is, as it were, in a new body. It is unclear to me that there is any reasonable standard to which to hold these two hypothesis. On the one hand, we have a case not unlike thought experiments on personal identity like Locke’s Prince and Cobbler, where all of a sudden one “person” finds himself in a completely different body (Call this “LPC”). That is essentially what would be happening on the hypothesis that D is conscious. On the other hand, we have the hypothesis that D is unconscious despite conscious behavior. It is certainly not clear that Kirk’s preferred hypothesis is more plausible. On LPC we do not have to break our standard habit of attributing certain mental states to an individual when he is in certain behavioral states. This should count as evidence in the favor of LPC. But, to me, there is no clearly preferable hypothesis.

In any case, let us return to the basic relevance of the Kirk example. Kirk has described a way that we can take a third-person perspective towards a zombie-like creature and have some positive evidence that “he” is a zombie. This would give us a way to satisfy Marcus’s worry if these sorts of behavioral anomalies were permitted in the zombie world. But it seems very bizarre to think that all the physical facts could be the same and behavioral facts could be different. As far as Chalmerian zombies go, there appears to be no such positive evidence for zombiehood. The reason we think zombies are unconscious, then, must be based on another sort of evidence. In sum, what Kirk gives us is a way to be justified in believing in zombies, even if we cannot positively conceive of their lack of experience. Kirk
does this by having Dan offer first-person reports of his diminishing consciousness. This helps a great deal in conceiving of zombies, but, again, cannot be used to help conceive of a zombie world, because it violates the constraint that zombies are our behavioral duplicates. There looks to be a genuine difficulty, as Marcus points out, in imagining zombies completely from a third-person point of view.

We might isolate the mechanism at work in Kirk’s example to see if there is a way, from a third-person stance, to imagine a zombie world. If we find such a mechanism, it would serve as a nice reply to Marcus. The mechanism at work in the case Kirk describes is that of verbal reports. Dan tells us that he is losing his consciousness piecemeal, despite the majority of his behavior suggesting otherwise. Might we be doing something like this in the case of the zombie world? Do we just need to find the analog in the zombie world in order to satisfy Marcus? One suggestion is that we just tell ourselves what Dan told us in Kirk’s case: that he (or the Chalmerian zombie) does not instantiate phenomenal properties. Let us take a closer look at this possibility.

**Intending, Supposition, and Stipulation in Imagination and Conceiving**

In this section I want to flesh out three related phenomena that are all extra-imagery factors that play roles in determining what it is that an image represents, which can be both non-actual and actual situations. The way I attempt to flesh out these three phenomena is by way of canvassing cases where two images are used to represent different situations and asking what makes the images represent different situations.

In *The Blue and Brown Books* Wittgenstein touches on the mechanism of intention in imagination,

---

What I mean by this locution should become clearer as the section progresses.
Someone says, he imagines King’s College on fire. We ask him: “How do you know that it’s King’s College you imagine on fire? Couldn’t it be a different building, very much like it? In fact, is your imagination so absolutely exact that there might not be a dozen buildings whose representation your image could be?”—And still you say: “There’s no doubt I imagine King’s College and no other building.” But can’t saying this be making the very connection we want? For saying it is like writing the words “Portrait of Mr. So-and-so” under a picture. It might have been that while you imagined King’s College on fire you said the words “King’s College is on fire”. But in very many cases you certainly don’t speak explanatory words in your mind while you have the image. And consider, even if you do, you are not going the whole way from your image to King’s College, but only to the words “King’s College”. The connection between these words and King’s College was, perhaps, made at another time. (Wittgenstein 1960: 39)

The connection was made, it seems, as one set out to imagine King’s College. There was most likely an intention to imagine King’s College that happened simultaneously with the start of the imagination process. The image by itself does not represent King’s College. One’s intention to imagine King’s College is what helps the image represent King’s College.¹

There is a closely related aspect of imagining that Peacocke (1985) brings out. Peacocke calls this “S-imagining.” The “S” stands for “supposition.” While this is not, as Peacocke notes, literally supposition, it is like supposition in that “it shares with supposition the property that what is S-imagined is not determined by the subject’s images” (Peacocke 1985: 25). The basic idea is that there is a mechanism that allows us to stipulate or suppose that certain background conditions are true in the situation we imagine. In what follows I will refer to this practice as “supposition.” I want “supposition” to have a rather narrow meaning here so that I am not interpreted as talking about another practice entirely, the

---

¹ I should note that I do not think there is a separation between intending to imagine X and actually imagining X. It would make no sense to say there is, first, an intention phase, which is followed by an imagination phase. The two happen at the same time. But for purposes of theoretical exposition, I speak, at times, as if they were discrete processes in order to examine them in their own right.
practice of supposing that such and such is the case. The following examples will throw these three related mechanisms, supposition, intention in imagination, and stipulation into relief.

**Some Case Studies: Same Image, Different Representation**

The first sort of case of supposition is given by way of a Cartesian example. I change things a bit in order to have two different products of imagination represent the same situation. In the *Meditations*, Descartes claims that while we can *imagine* a pentagon, we cannot *imagine* a chiliagon. However, we can *conceive* of a chiliagon.

Let us suppose that we have a cerebroscope hooked up to someone reading this section of Descartes’s *Meditations*. This device allows us to examine the reader’s mental images by capturing the phenomenology of his mental states and then transfer these images into our heads, allowing us to share the images. Suppose that the images the reader gets during the passage about the chiliagon and pentagon are identical. The subject tells us that he is imagining, first, a pentagon and, second, a chiliagon. We “see” them as the same (remember we get his images), but we know from his report that they represent different things. What mechanism is at work that distinguishes one from the other?

Descartes thinks this differentiating mechanism belongs to the intellectual faculty of conceiving. What we turn to when we conceive of chiliagon is not an image, e.g. the image of a pentagon (or the image of a near-circle), but to the very concept or definition of a

---

10 The general practice is very close to what Chalmers’s calls “negative conceivability”. When I use “supposition,” though, I just mean to refer to the extra-imagery work we do in representing situations by way of imagination.

11 In the rest of the essay, I employ the notion of representation to discuss imagination. I refer to the *product* of the *process* of imagination as a representation. This product represents, again, a way the world is or could be. This product will often include more than just an image. Many times it will also include the extra-imagery features that this section is meant to flesh out.
chiliagon. We appeal to something outside of the image. We can still speak of the pentagonal image as *part of* a proper representation of a chiliagon, but there is clearly something else, aside from mere pictorial correspondence, at work. We might say that what is doing the work here is a supposition we make about the representation and our concept of a chiliagon such that we make one image represent a chiliagon and the other not. Of course, it is not just an arbitrary stipulation that whatever is in my mind can correctly represent anything.\(^\text{12}\)

What happens in the chiliagon case is that my ability to picture or visualize a 1000-sided figure fails at a certain point. I get a faint image that is roughly circular and then, in virtue of either **supposing that it is a chiliagon** or **intending to represent a chiliagon**\(^\text{13}\) and understanding that there is no contradiction in such a figure (an instance of supposition) I succeed in imagining it.

Consider another case. When I imagine the actual world and the zombie world, down to each microphysical detail, there is no distinction in the two mental representations at the level of images. So, there must be something “outside” of the image that makes one a correct representation of the actual and the other a correct representation of the zombie world.\(^\text{14}\) What makes one a representation of the zombie world and the other not could be one of two things. First we might just intend to imagine a zombie world. Second, someone may ask us to imagine a world that is identical to the actual world in all respects and then

\(^\text{12}\) For example, I cannot just image a red chair and then suppose that this image represents a world in which the laws of nature are different from the actual world. More would be required in this case.

\(^\text{13}\) The way I see it, intention is a mechanism that happens before we set out to imagine, and supposition is a mechanism that we can apply to an image we already have. Supposition is like stipulation. But I use “stipulation” to refer to conditions that are stipulated by a third party and not by the imaginer.

\(^\text{14}\) Recall that Marcus seems to think that since the distinction is at the level of phenomenal properties, we must imagine the absence of these in the zombie world. But we have no way to do this. And since we cannot imagine subjective absence, Marcus claims zombies are (positively) inconceivable.
asks us whether we can subtract experience from that world and still maintain a coherent representation. If we can, it seems that we achieve this result by doing something like what we did in the chiliagon case: we turn to the concept of consciousness and attempt to detect a contradiction in a world like ours without experience. In a word, we are supposing that certain conditions hold in the imagined situation. Given that image, can we coherently suppose that such and such to be the case in that world without reaching a contradiction?  

It appears that the role of intending in imagination mirrors the role intending plays in pictorial representation. When I draw a picture of Descartes on the blackboard it does not by itself represent Descartes. It needs to be drawn under the right sort of conditions: I need to intend to represent Descartes as I set out to draw the stick figure. I can make my intention vivid by writing “Descartes” under the stick figure. These sorts of things, both the intention to achieve a goal, and the stipulation (writing “Descartes”)—a sort of intention—are what make my representation a representation of Descartes.

More Case Studies: Imagining or Conceiving of Absence

There are other sorts of cases that help to show the product of imagination is often more than an image. These cases—where we are asked to imagine a world without X—are particularly relevant to the zombie case. Let us not worry about zombies right now though.

---

15 Here is another sort of case. Consider your favorite set of identical twins. Imagine just one of them. What is it about your representation that makes it about that twin and not the other one? It seems to be the fact that we have a certain goal in mind, a certain intention, and then draw up a relevant image, and then we see that we have achieved the goal.

16 This is not due to any fundamental distinction between images and pictures. We shall see later that images are less like pictures than we might intuitively think.

17 It is not difficult to represent Descartes in this way. Similarly it is very easy to imagine Descartes in his study, despite never actually seeing it. The reason is because something similar is happening in both cases: there is the same tight connection between goal and supposition.
Instead, consider a world identical to ours in terms of appearance properties but with no abstract objects. There will be tokens but no types. What makes this world different from ours, supposing that abstract objects exist in our world? What is it about the product of the imaginative process that makes it correctly represent an abstracta-less world?\(^{18}\)

It seems that in this case it is both our intention to imagine an abstracta-less world and our ability to coherently suppose certain conditions to hold in the imagined world in question. We can “see” that there are math books filled with numerals but—we suppose that certain conditions hold—there are no numbers. Since this supposition does not entail a contradiction, at least does not seem to, we can *prima facie* conceive of a world just like ours in terms of appearance properties but with no abstracta. Of course, upon closer reflection there may be a conceptual contradiction lurking in the hypothesis of a world with concreta but abstracta, but that would not be obvious.

Stipulation also plays a role in imagination.\(^{19}\) It appears that in a quote mentioned earlier Chalmers is doing something close to stipulating that zombies lack consciousness. For what (positive) reason do we have to doubt zombies’ first-person reports of their own conscious states? Chalmers stipulates the there is nothing it is like to be a zombie, but how can we guarantee this lack of phenomenal properties? Let me just dwell on this a bit, since the point I want to make here is one of our major threads.

To see why supposition and its third-person analog, stipulation, are such powerful mechanisms in imagination, we need to look at a feature of imagination that we might call the *prima facie* inability to misrepresent what one is imagining. Let me explain what I mean

---

\(^{18}\) Ignore the move that says that it is clearly a representation of the actual world, and the actual world is a world with no abstracta.

\(^{19}\) Kripke makes this plain (1980: 44). He says that possible worlds are stipulated not discovered. So stipulation plays a real role in determining that which we are to imagine.
by this. When I ask you to imagine your father reading the newspaper, before you even form the image, there is no question that what you will imagine will be about your father. The goal (given by my stipulation) to imagine your father heavily influences the content of the product of your imaginative process.

When I ask you to imagine a zombie world, I stipulate that zombies are not conscious, so, whatever it is you imagine, there is a huge presumption in favor of your product seeming to represent the zombie world. Oftentimes in imagination we find ourselves under the influence of stipulation, which works along the lines of a self-fulfilling prophecy. We get a description of a situation where it is stipulated that a creature in question lacks normal p-consciousness and take it for granted that the situation as described by the stipulation is coherent. Recall Chalmers’s stipulation,

At most this shows that from the third-person point of view, my zombie twin and I are identical, so that you cannot be certain that I am conscious; but we know this all along. But it does nothing to imply that from the first-person point of view, I cannot know that I am conscious. From the first-person point of view, my zombie twin and I are very different: I have experiences, and he does not. (Chalmers 1996: 199)

In the last sentence, Chalmers clearly just stipulates that our zombie twins are unconscious. But how else could one show that one’s zombie twin is not conscious? It seems that we have to stipulate that they are unconscious, but how can we be sure that they are? I discuss this question later. For now we can note that Kirk’s evidence for zombies is clearly preferable to stipulation or supposition—but there appears to be no such evidence in the zombie world.

---

20 I want to use “stipulation” to refer to a condition imposed on the product of imagination by a third party, and use “supposition” to be a mechanism that the “first-person” uses to impose certain conditions on the product of their imagination.
Searle’s Silicon Brain and the First-Person Perspective in Imagination

If Robert Kirk gives us a plausible way to imagine zombies from the third-person perspective, then John Searle (1992) offers what seems to be a plausible first-person way. The main reason to discuss this case is that it seems to provide a case of imagining zombies from the first-person perspective. If we want to be sure that the creatures we imagine lack consciousness, what better way than to locate our own consciousness in the world and then slowly remove it?

By way of a series of thought experiments, Searle tries to show that our mental concepts are fundamentally distinct from behavioral concepts. In these thought experiments one is asked to imagine his neurons being replaced with silicon chips, although one could run the thought experiment without the replacement. In any case, this replacement can lead to three different outcomes: (1) no functional or experiential difference, (2) experiential difference without functional difference, (3) functional difference without experiential difference. Since (2) is close to the zombie scenario, we shall dwell on it. Searle describes this possibility as follows:

as the silicon is progressively implanted into your dwindling brain, you find that the area of your conscious experience is shrinking, but that this shows no effect on your external behavior. You find, to your total amazement, that you are indeed losing control of your external behavior. You find, for example, that when the doctors test your vision, you hear them say, “We are holding up a red object in front of you; please tell us what you see.” You want to cry out, “I can’t see anything. I’m going totally blind.” But you hear your voice saying in a way that is completely out of your control, “I see a red object in front of me.” If we carry this thought experiment out to the limit, we get a much more depressing result than last time. We imagine that your conscious experience slowly shrinks to nothing, while your externally observable behavior remains the same. (Searle 1992: 66-67)

While in this scenario you are not a zombie in the beginning, your phenomenal experience is different from the normal way we experience the world, with our mental states causing our behavior. For example, when I desire to get up to get a cup of coffee, I expect
the right sort of behavior to follow. In fact, I operate under the assumption that after this initial decision, my mental states continue to cause the majority of the behavior that gets me from my desk to the coffee maker. In the thought experiment, my experience is not like that. I am conscious but not in a way that my conscious mental states cause behavior. My behavior is caused by the functional states instantiated by certain patterns of silicon chips.

When we imagine such a thing happening, there does not seem to be any incoherence. This suggests that we have a way to meet Marcus’s demand that we imagine subjective absence, provided, of course, we can coherently think of the being that was us as unconscious. The thought experiment gets modified a bit, because on Searle’s rendition the world is our complete duplicate at one time, and then later it is a duplicate of our world save consciousness. So, we know that at time $t_1$, we are conscious. But we can coherently imagine that the experience gets progressively weaker until it disappears altogether at time $t_2$. Is this enough to ease Marcus’s worry?

After all, in the Searle case, we can imagine subjective presence from the first-person perspective. Then we imagine subjective absence from the third-person perspective. This move seems to give us a way to pull off the trick. But I doubt Marcus would be satisfied, for the problem is just reinstated: We still have no way to determine that the body that was ours is not now conscious. Perhaps it is inhabited by a different person. The new “you” talks and walks as if he were conscious. It seems to be conscious. What grounds have we for thinking that he lacks consciousness? It seems that we have even less to go on than in Kirk’s case. We seem, in fact, back to just supposing that he lacks consciousness.

**Zombie Dave**

In *The Conscious Mind* David Chalmers writes that there is no great difficulty in conceiving of zombies, understanding conceivability as the conceptual coherence of a given
situation (Chalmers 1996). Since I want to pay particular attention to how we are to think of
zombies or the zombie world, so we shall look at what Chalmers says.

So let us consider my zombie twin. This creature is molecule for molecule identical to
me, and identical in all the low-level properties postulated by a completed physics, but
he lacks conscious experience entirely. (Chalmers 1996: 94)

I confess that the logical possibility of zombies seems equally obvious to me [as that of
a mile-high unicycle]. A zombie is just something physically identical to me, but which
has no conscious experience—all is dark inside. While this is probably empirically
impossible, it certainly seems that a coherent situation is described; I can discern no
contradiction in the description. In some ways an assertion of this logical possibility
comes down to a brute intuition, but no more so than with the unicycle. Almost
everybody, it seems to me, is capable of conceiving of this possibility. (Chalmers 1996:
96)

What I want to bring out in these examples is the type of conceivability that Chalmers
has in mind. Chalmers thinks most of us, who are not in the grips of a theory (e.g., Analytic
Functionalism), can imagine zombies.

How? Well one way seems to be the same way that we conceive of a mile-high
unicycle, which is by examining the intensions (meanings) of the relevant concepts
involved.21 If we have a proper understanding of UNICYCLE22, then we can see there is not
anything about UNICYCLE that rules out the possibility that the objects that fall under it be
ridiculously tall, and vice versa for MILE-HIGH. Note that certain things can be ruled out: a
mile-high number, for example. If we are competent uses of MILE-HIGH then we
understand that it can apply (non-metaphorically, that is) only to physical objects.

Applying this method to the case of zombies, we look, according to Chalmers to the
following concepts, PHYSICAL STATE, FUNCTIONAL STATE, BIOLOGICAL

21 The appeal to intensions may give some physicalists pause right off. I do not think much
hangs on this though, since I suppose there is a way to argue that intensions are
ultimately mental states, which the physicalist can say are reducible to physical states.

22 Capitalized words and phrases denote concepts.
ORGANIZATION, ORGANISM, etcetera. Chalmers thinks that there is nothing in these concepts that requires (conceptually) that there be something it is like to be a creature that has the properties that fall under these concepts.

A frustrating aspect of this method is that it is unclear how we are to examine the content of the relevant concepts. Presumably we have to conceive of a situation that has properties that fall under those concepts. Chalmers does not say this, though. The only way I can test the coherence of a conjunction of these concepts is by directing myself to a situation that requires me to use all the relevant concepts. It seems to me that in almost all cases the only way to test the conceptual coherence of a hypothesis is by conceiving of a situation that has all the right sorts of properties, properties that fall under the right concepts. I do not have a clear idea of how to conceive in the absence of a situation, so it is a bit misleading on Chalmers’s part to think that zombies are as simple as mile-high unicycles, because the mile-high unicycle case gives us the impression that all we need do is compare intensions. I do not see how such a pure-intension-checking feat is possible when it comes to zombies. It seems to me that we can test the coherence only by imagining/conceiving of a situation in which all the concepts get deployed.

Chalmers does offer some help: we can imagine zombie by thinking about the actual world. He writes,

To fix ideas, we can imagine that right now I am gazing out the window, experiencing some nice green sensations from seeing the trees outside, having pleasant taste experiences through munching on a chocolate bar, and feeling a dull aching sensation in my right shoulder. What is going on in my zombie twin? He is physically identical to me, and we may as well suppose that he is embedded in an identical environment. He will certainly be identical to me functionally: he will be processing the same sorts of information, reacting in a similar way to inputs, with his internal configurations being modified appropriately and with indistinguishable results. He will be psychologically

---

23 I cannot see how it would help to examine the concepts independently of such a situation, at least not in the case of zombies.
identical to me, in the sense developed in Chapter 1. He will be perceiving the trees outside, in the functional sense, and tasting the chocolate, in the psychologically sense. All of this follows logically from the fact that he is physically identical to me, by virtue of the functional analysis of psychological notions. He will even be “conscious” in the functional sense described earlier—he will be awake, able to report the contents of his internal states, able to focus attention in various places, and so on. It is just that none of this functioning will be accompanied by any real conscious experience. There will be no phenomenal feel. There will be nothing it is like to be a zombie. (Chalmers 1996: 95)

This example requires that we “fix” the phenomenal facts by imagining our own qualitative feels, and then imagine ourselves (or someone looking like us) from a third person perspective. Of course, we are not imagining ourselves being zombies. But if we are merely imagining beings that look like us from a third-person view, then it is not clear that we are conceiving of zombies, for Marcus’s reasons.

Summary

Before moving on to the sketch of a theory of imagination, I think it is important to note that many admissions of the zombie intuition are accompanied with an interesting caveat. A good representation is in Frank Jackson’s “The Case for A Priori Physicalism” where he says that the zombie world “seems offhand to be possible” (Jackson: 259). Most philosophers claim that zombies are prima facie conceivable. This is an important admission because it leaves it open whether the conceiving is ideal.

Let us briefly review some of the findings of this chapter. The imitation man section raised doubts about whether we can coherently conceive of non-phenomenal occurrent beliefs and tried to show that creatures can lack sensory-modality-based phenomenology but still instantiate phenomenal properties. Block’s blindsighter fleshed out a number of issues: the distinction between a-consciousness and p-consciousness and the extent to which we can conceive of them coming apart. Siewert’s cases helped address skepticism about splicing and perspectival imagination. In the Kirk section we discussed the role of positive evidence and
the role of intention, supposition, and stipulation, and the phenomenon of appealing to the
latter two in the absence of positive evidence. The Searle case seemed to offer a way to
imagine first-person absence, but it has complications that make such a method problematic.
Last we look at Chalmers’s instructions for imagining zombies. We noted two things here.
First we noted that his method of checking intensions is at best just elliptical for talking
about situations, second, that the way Chalmers describes zombies seems to fall prey to Eric
Marcus’s worry that we have no good evidence for thinking that zombies lack p-
consciousness.
CHAPTER 3
A SKETCH OF A THEORY OF IMAGINATION/CONCEIVING

In Chapter 2 we examined imagination by looking at thought experiments in the
philosophy of mind. The hope was that we would see that we should proceed with more
care based on individual worries. This section picks up that thread, but the work here is of a
more general nature. Here I offer a sketch of the structure of imagination and conceiving.

Imagination is a pervasive phenomenon inside and outside the philosophical study.
We lose ourselves in daydreams. We find out thoughts drift to dinner during uninspired
lectures and tedious conversations. Much of our mental lives are spent representing a
“world” distinct from the one in front of us. Philosophers spend time imagining
hypothetical situations to test coherence (Zombies) or to illicit intuitions about what we
might say of such imaginary situations (Twin Earth). Philosophers in search of the nature of
various entities ask us to conceive of situations—a clear dependence on the mechanism of
imagination—but rarely is the lens of examination turned on this mechanism itself.

In searching for the nature of imagination, we see a rough model in our ordinary
imagination; in philosophy and ordinary life we *represent* various ways the world was or could
be. To *represent* a situation (to oneself) is the essence of imagination. To say that imagination
represents things is to say that it is intentional in the philosophical sense of being about
something. ¹ It is also intentional in the everyday sense in that it is purposeful.² These are two
important facts on which to frame our discussion of imagination.

¹Searle makes this point (Searle 1983: 4)
²McGinn (2004) and Casey (1978) make the point.
Situations

Descriptions of imagination reveal an interesting structural fact. Consider the following sentence:

Jones imagines George Bush working in the oval office.

What follows “imagines” in this case seem best described as “a situation” or “a state of affairs.” It need not be an action, as in the case above, it could be an event with very little going on. I could ask you to imagine a table. It seems that no matter what it is that follows ‘imagine’ it will be a situation. Even in descriptions like

Jones imagines George Bush

it seems that there will still be a situation that Jones has in mind. If one tries to imagine Bush simpliciter, he will find that even in cases where he just imagines a smirking face; that face is still a situation, even if it is a minimal one.

Descriptions of conceiving have a different grammatical feature than descriptions of imagination, namely the “of” that follows “conceiving” is missing in descriptions of imagination. We are usually said to imagine \( x \), while we conceive of \( x \). Even so, when we conceive, we almost always conceive of a situation. It is ungrammatical to say,

Jones conceives George Bush.

---

3 I speak of imagination here and leave out conceiving because I think that what many have in mind when they think of conceiving is just imagination of a special sort.

4 I use “situation” here, but I could have used “state of affairs” as I cannot see any difference between the sorts of things these locutions pick out.

5 If we allow Chalmers’s negative conceivability (below) to be counted as conceivability, then it looks like some cases of conceiving will not be cases of conceiving of a situation, but of reasoning in such a way that we cannot rule something out.
We can, of course, substitute “imagines” for “conceives” here and have a grammatical construction. Yet I do not think this shows a deep distinction between conceiving and imagination. Rather, I think it points to their deep connection: for, as in imagination, Jones must conceive of George Bush in relation to something else. I cannot conceive of Bush simpliciter.7

Consider the following sentences that a subject might use to describe an occurrent mental state:

1. I think that the president of the United States is Al Gore.
2. I imagine the president of the United States being Al Gore.
3. I conceive of the president of the United States being Al Gore.
4. I entertain the thought that the president of the United States is Al Gore.

It seems that one who genuinely asserts (2) and (3) will be in a phenomenologically richer mental state than when he is undergoing the acts he reports in asserting (1) and (4).

(4) is an attitude, at least on a common understanding of ‘entertain’, that seems less like imagination and more like thought or pure conceiving in terms of phenomenology. In terms of the nature of the attitude one takes to the that-clause it seems closer to imagination and conceiving because one does not express a belief when they utter (2), (3), or (4) assertorically as one does in uttering (1). For now, let us say that conceiving, entertaining,

---

6 Joe Levine gives a similar definition of a situation. He writes, “By a situation I mean an object’s instantiating one or more properties (perhaps at a time), or an ordered n-tuple of objects instantiating one or more relations.” (Levine 2001: 40)

7 I may, however, be able to think about Bush simpliciter by way of just reading a sentence, part of which refers to him. For example, suppose that I am reading about a trip the president took to France in the newspaper. It seems that I can comprehend what I am reading without having to imagine him in France. I take this phenomenological fact to support the claim that I cannot conceive of or imagine just Bush in the way I seem to be able to just think about him.
and imagining are propositional attitudes that differ from belief in that they do not require that the subject decide on the truth of the proposition\(^8\) to which the attitude relates.

**Target Situations and the Subject’s Description**

Let us introduce a bit of terminology here to help sketch a bit more of the structure. Let us call what follows “imagine” in sentences of the sort listed above the “target situation,” and let us call the way that the subject describes the target situation “the subject’s description” of the target situation. Target situations, or the goal of the imaginative act, can be generated from a variety of sources. The generative source of a target situation may be a philosophy paper. One might try to imagine what it is like to be a bat when one reads Nagel. If so, *what it is like to be a bat* is the target situation. One might find himself imagining a strange situation as a result of a conversation one is having. The target situation is that which the subject tries to imagine. A subject’s description is how he describes what he imagines.\(^9\)

**Propositions, Content and Representations\(^10\)**

The subject’s description of his representation of the target situation is a way to articulate what philosophers call “propositional content” or “intentional content”. In the first Bush case, the target situation is *George Bush working in his office*. And let us say that the subject’s description is *George Bush pushing papers around on his desk to look busy*. The

---

\(^8\) I say that these are propositional attitudes, which may sound strange since I just said that situations are what one is related to in conceiving and imagining. But I make the connection between propositions and situations plain later. In order to make a first pass I use ‘proposition’ and leave it at that for ease of exposition.

\(^9\) Bernard Williams develops a similar notion, which he calls “narration” (Strawson 1969: 197). He also makes the point, along with Peacocke, that a subject can be mistaken about that which they imagine.

\(^10\) Peacocke calls these “imaginations” (Peacocke 1985). He uses this phrase to refer to not just the image but the image plus that which is outside of it. In my terminology, Peacocke’s use of ‘imaginations’ is synonymous with the product of the process of imagination.
propositional content, then, is *that George Bush is pushing papers around his desk to look busy*. At this time we need to note that there is a sense in which imaginations are infallible and a sense in which they are not.¹¹ The subject’s description is infallible because there is no way that that he can be mistaken about what he *seems* to be imagining. But we surely do not want to say that he is always the final authority on the success of his imagination. He may be mistaken about whether the product of his imaginative process correctly represents the target situation. As Christopher Peacocke warns, “we have to distinguish sharply between the intention to imagine something and the success of that intention” (Peacocke 1985: 26).

While I use the phrase “propositional content” I do not mean to suggest that when one imagines one is directed at a proposition. What we are directed at is a situation. It should be a bit unclear how we go from talking of situations to propositions, so let me try to make that relationship less murky. In *Mindsight* Colin McGinn claims that what he calls cognitive imagination, the sort of thing philosophers do, is a propositional attitude. Imagining, like believing, desiring, wanting, McGinn thinks, is an attitude one takes to a proposition. This way of talking comports with my view of talking about situations, because when one looks at it, it is easy to see that one is not directed at a proposition in all acts of imagining and conceiving. Talking about being directed at a proposition is only shorthand for being directed at a situation that is accurately described by that proposition. For example, most of us would not want to say that one *imagines* a world where everything is red just by way of being directed at the proposition *that everything in the world is red*. Some might say that sort of relation is sufficient for thought (provided he is willing to countenance mind-independent

---

¹¹ Peacocke notes this. He writes, “It is sometimes said that the reason it is absurd to suppose that an imaginer might be mistaken about what he is imagining is that imagining is done intentionally, and in general we have non-inferential knowledge of our intentions” (Peacocke 1985: 26).
propositions)—but surely it is not sufficient for imagination. Rather, in imagination and
conceiving, I imagine a situation in which things are that way, and then by way of describing
the imagination I “give” the propositional content that everything in the world is red. So on my
view propositional content is a function of the product of the imaginative process and
subject’s description, which need not be verbal.

This is related to a distinction that Stephen Yablo makes between what he calls
objectual and propositional imagining (1993). He thinks that propositional imagining has
alethic content and objectual imagining has referential content. Yablo says that alethic
content is the sort that can be true or false, while referential content is not: it can only refer.
An alethic content bearing mental state is true or false of something, while a referential
content bearing mental state refers or does not. Let us take a look at Yablo’s example to fix
ideas,

Imagining can be either propositional—imagining that there is a tiger behind the
curtain—or objectual—imagining the tiger itself. (Yablo 1993: 27)

In both cases, the image may be the same, but there is something different going on in one
of the imaginations such that it is true or false because it expresses a proposition. It is a bit
unclear on what this difference-making mechanism is. What does Yablo say?

To be sure, in imagining the tiger, I imagine it endowed with certain properties. So
objectual imagining has in some cases propositional accompaniment. Still the two
kinds of imagining are distinct, for only the second has alethic content—the kind that
can be evaluated as true or false—and only the first has referential content—the kind
that purports to depict an object. (Yablo 1993: 27)

What expresses a proposition is the subject’s description of the imagination or a supposition
one makes about the image: as in Yablo’s case of supposing it is endowed with certain
properties. This need not be verbal of course, as Wittgenstein reminds us, “in very many
cases you certainly don’t speak explanatory words in your mind while you have the image”
To make this clear—this phenomenon of a subject expressing the propositional content of his imagination by taking a “look” at his image, so to speak—let us look to an analogy. Let us ask what it is about sentences that make them express different propositions. Consider the following sentence:

John lies outside the house.

Used assertorically this sentence can express (at least) two propositions, which could also be expressed, respectively, by the following sentences,

John is prone on the ground in an area sufficiently near but not in the house.

John tells falsehoods with the intent to deceive in an area sufficiently near but not in the house.

That is, one could assert “John lies outside the house” on different occasions and express two different propositions in each occasion. What makes the difference? After all, the sentence is the same is it not? One possible candidate here is the asserter’s intentions. After hearing him speak, we still have to ask (if it is not clear from context): What was he trying to say? Sentences do not express anything in and of themselves. We use them to express thoughts, but they do not do all the work. Similarly, images do not represent things in and of themselves. They can represent things only if there is an intention (or supposition) on the subject’s part to imagine such and such.

If that is a plausible account, then we might think that one’s intentions to imagine such-and-such might do the work of intentions in the case of uttering sentences to express propositions. So it is my image plus my intention to imagine such-and-such that gives the propositional content. This does not exclude the possibility that I, after first intending to imagine such-and-such, can still suppose that certain things are true in the imagined world
and change the product from what I initially intended. So intention or supposition can do
this work.

Now, with those observations on the table, a curious link emerges. Given the role that
intention plays in imagination and the influence that the target situation has on one’s
intentions, we can see that one’s intention to imagine such-and-such is heavily influenced by
how the case is stipulated (the target situation that is). In fact, in the majority of cases, the
subject’s description of what he imagines will be, in so many words the target situation.12

Representations and “Verification”

Now for the question of what sense it makes to say that the representations we have
when we imagine and conceive are truth-evaluable. We need to get clear on what makes
“Zombies are conceivable” true. Answering this question will also help us make more sense
of the notion of verification that Chalmers deploys in his important discussion of conceiving
(2002). “We can say,” Chalmers writes, “that an imagined situation verifies S when reflection
on the situation reveals it is a situation in which S” (Chalmers 2002: 152). “S” stands for
statement, which for Chalmers is a sentence with truth conditions. I think this way of setting
things up gives us a way to understand when and how the products of imagination correctly

12 There are cases where we imaginatively represent something and that representation lacks
propositional content. For an example, let us think of the tiger example. One might form a
mental representation while reading a book about tigers. Suppose one is engrossed in the
book and then suddenly sees the curtain move out of the corner of his eye. If he is locked in
deep thought about tigers, it is not unbelievable that he would form a representation of a
tiger behind the curtain, out of the blue, so to speak—he has tigers on the brain after all. But
according to the story I am telling that merely objectual representation, which he gets with
an atypical target situation, itself, will not be sufficient to get the propositional content that a
tiger is behind the curtain. That representation would have what Yablo calls referential content,
but in order for propositional content something more is needed. The best I can say about
the something more is the sort of supposition one makes about what one is imagining or
previous intention to imagine such-and-such (which is clearly missing in the random tiger
image case just mentioned).
represent the target situations. One correctly represents a target situation when one’s product has content that matches the target situation. When this happens we can say that a target situation is verified by the content or that the content accurately represents the target situation. Zombies are conceivable just in case the content of the product of our imaginative process correctly represents the relevant target situation (the third-person stipulation of the zombie world). This is quite similar to Chalmers’s view (2002), except I talk of the content matching the target situation instead of the statement being verified by the imagined situation. The details are different, but I think the overall pictures are the same.

But is this notion of verification the sort of relationship that allows us to say whether a product of an imaginative process is true or false? The claim “x is conceivable” is true relative to a subject just in case the subject’s product of imagination plus his description of the product—which need not be verbal—“gives” propositional content that matches up with, or verifies, the target situation.¹³

So as long as the subject finds the product of imagination coherent and to correctly represent the target situation, then he will find that it is capable of being an accurate representation of the target situation. The content of the imaginative product will match up correctly or correct represent the target situation. For example, if I am asked to imagine a mile high unicycle, first I have a representation of a situation. Then I decide whether my imagination is coherent and in doing so express the propositional content that there is a mile

---

¹³ Now we can also say that the subject’s description is true provided it correctly describes the product of imagination. On my view, this is a requirement for expressing the propositional content that enters into the verification relationship with the target situation.
high unicycle. This content is true provided I find it correctly verifies the target situation. It will typically do so, provided I am not suffering from a confusion.

Now this reveals a curious feature of imagining and conceiving. The truth makers of the products of imaginative processes are mind-dependent. On this view what makes a representation in imagination correct or incorrect is what is in the subject’s head and his judgment of whether or not it coherently depicts or describes the target situation and facts about whether or not he is confused. Couple this with the presumption in favor of the subject’s description being, in so many words, the target situation, and we can see that there is a presumption in favor of success. The structure of imagination has the form of self-fulfilling prophecy.

Subject’s Descriptions and Fallibility

On my view, imagining is an intentional state, a state that represents something. This may naturally lead one to infer that the target situation is what the products of imagination are about and that the subject’s description always correctly describes that which, in Peacocke’s words, is “the causal source of the image” (Peacocke 1985). Intuitively, we might think that we have a certain authority over the content of our mental states in general. Why would the mental states of imagination be any different? To motivate this infallibility a bit, consider the following quote from Wittgenstein:

If a man says “I am imagining King’s College on fire,” it seems absurd for his friend to respond ‘How do you know it is King’s College you are imagining? Are you aware that there is a precise replica of it on a film set in Los Angeles? May you not be imagining it? (Wittgenstein 1965: 39)

This quote seems to suggest that subject’s cannot be mistaken about that which they imagine. It seems that the sole determiner of the representation’s accuracy is the subject’s intention that it represent King’s College (or supposing that what he imagines to be on fire is King’s College).
This is not to say, however, that this makes it an accurate representation. In one sense, the subject cannot be wrong: he cannot be wrong about what the image seems like to him. He can, however, be mistaken about the causal source of the image. Peacocke articulates this thought:

Subjects can easily be mistaken about which object their image is causally of, there is no absurdity in asking questions which presuppose that they might be mistaken.

(Peacocke 1985: 27)

The sense in which it is absurd to ask the imaginer whether his image is of King’s College is the sense in which we ask what it seems to him that his image represents. He cannot be mistaken about this. However, whether the causal source of his image (the actual King’s College) is correctly represented by his imagination is another question. Perhaps he has been duped by a playful don into thinking that Baliol College is King’s College. This would be a scenario in which he mis-imagines King’s College because what he actual imagines is Baliol College.

Later in the *Philosophical Investigations* we find Wittgenstein struggling to put his finger on this phenomenon. He writes,

I draw a head. You ask “Whom is that supposed to represent?”—I: “It’s supposed to be N.”—You: “But it doesn’t look like him; if anything, it’s rather like M.”—When I said it represented N.—was I establishing a connexion or reporting one? And what connexion did exist? (Section 683)

It seems that if supposition is as I describe it, then it is reporting a connection, provided that when Wittgenstein draws the head, he intends it to be of N. If he just draws the head first with no intention to draw anyone in particular and then decrees, after the question, that it is N, then his words report a connection.

But despite the seeming infallibility, we can be mistaken about what it is we are imagining. I cannot be mistaken about what I take myself to be imagining. If it seems to me that I am imagining Bill Clinton, then it would be absurd to say that I can be mistaken about
the seeming. But if I am under the misapprehension that Jimmy Carter=Bill Clinton, and my
image is causally of Jimmy Carter and not Bill Clinton, then despite my intention to and
conviction that I imagine Bill Clinton, it seems obvious that I fail to do so. Thus we can see
a subject’s descriptions are fallible. Now let us turn to other cases where we see the so-called
infallibility of a subject’s description break down.

**Fallacies of Imagination**

**The Illusion of Immanence**

Many times when we imagine a situation, we turn it over in our minds. If I ask you to
imagine George Bush in his office, you get an image in your mind. Now when I ask you to
imagine George Bush in his golf cart you get another image. Then if I ask you to imagine
Bush in the oval office again it seems as if you turn right back to the very same image you
just had. It seems as if you are directed right at it.

When one performs exercises like this, it is tempting to think that there is a stable
image in one’s mind that is an intentional object we can turn back to in order to extract more
detail of the image. This is a natural, but ultimately confused, picture of how imagination
works. It is a confusion of process and product or act and object. Jean-Paul Sartre warned
against this, labeling it the “illusion of immanence” (1963). It is the fallacy of mistaking the
image for the intentional object. Caused by the vivid nature of images and easy with which
we call them up, the illusion of immanence, in Sartre’s words,

assumes implicitly the existence of two complementary worlds, one consisting of
things, the other of images, and that when one of them becomes faint the other is
thereby illuminated. This is putting images on the same plane with things, giving both
the same sort of existence . . . paying attention to these images themselves must be
carefully avoided. (Sartre 1963: 61)
Images are not objects in their own right; they are of objects (and properties). Careful attention to the phenomenology of imagining shows why it is tempting to think otherwise. You yourself may have done so in the brief exercise above. One mistakes the previous image one had for the intentional object, which is to say one takes oneself to be directed right at the image. I think the common idea here is that there are pre-fabricated images in the mind that we can draw up at will. But this is a mistake, as the following example shows. Imagine your living room. Now try to focus only on the image of your living room. You cannot, as you will find, sustain the image of your living room without returning to your actual living room. The phenomenology here suggests that we need to turn the mind back to the actual world to keep the image before our minds.\textsuperscript{14,15}

\textbf{Is There an Illusion of Immanence?}

Perhaps you are unconvinced by this so-called illusion. Let us look at two arguments in support of the claim that images are not intentional objects. In \textit{Mindsight}, Colin McGinn
\begin{itemize}
\item[\textsuperscript{14}] While one is at it, it might not be a bad idea to note a bit more phenomenology here. It seems that what I’m calling the image of one’s living room is not all that image-like, at least not in the sense that an image closely resembles a picture. It does not seem like a picture. When one reflects on it, it seems that what we have been calling ‘the image’ is actually not all that imagistic. There are a few vague points of reference to your actual living room—a chair here, a couch there in the vaguest outline—but nothing that has any great detail, at least not to me.
\item[\textsuperscript{15}] Sartre also makes the point another way. His maxim is that consciousness is consciousness of something. We cannot imagine such-and-such if our image is not of such-and-such. We are simply persuaded by deceptive phenomenology here:
\begin{quote}
An imaginative consciousness is consciousness of an object as an image and not consciousness of an image. But if we form a second consciousness, or a reflective consciousness, on top of this imaginative consciousness, a second kind of belief appears: the belief in the existence of the image. (Sartre 1963: 125)
\end{quote}
\end{itemize}

The way in which we imagine is so quick that we often fail to grasp the intricate process by which objects as images appear to us.
(2004) presents two arguments designed to show that the illusion of immanence is just that. He targets “the Picture Theory” (or Sense-datum Theory). On this view we are aware of only our own mental images (or sense-data). These mental images represent the mind-independent world as being a certain way, but we never directly imagine the external world. We form pictures of the world that have all the properties we attribute to external objects. These “pictures” are epistemic intermediaries. Or, in our intentional-object talk, images are direct intentional objects which represent mind-independent objects, the indirect intentional objects.

The first argument against this picture is called the “Medium Objection.” It goes like this. If the image is a picture, it should have some non-intentional properties as all pictures do, and so, we should be able to turn our attention directly to those, irrespective of the direct object. The living room case I gave earlier is a version of this argument. Performing this exercise makes images look less like pictures since the exercise requires us to go back to the living room (which McGinn thinks is the direct object). McGinn writes, “I cannot turn my attention to the materials of the image independently of what it is an image of” (McGinn 2004: 63). This is essentially a version of Gilbert Harman’s argument in “The Intrinsic Quality of Experience” that appeals to the transparency of experience to show that the idea that vision is mediated by mind-dependent intermediaries is false, except it is given in the context of imagination instead of perception (Block, Flanagan Güzeldere 1997: 663-675).

The conclusion of the second argument is that the Picture Theory leads to an infinite regress of “pictures,” since the theory, according to McGinn, is committed to the general principle that awareness of an object must always be mediated by awareness of a sense-data. So, it seems the picture-theorist is committed to the view that we can be aware of any given sense-datum only by having another sense-datum, since sense-data are objects. If this
argument is sound, it is a real problem not only because we do not attend to such a multitude of “pictures” in imagination, but also because a regress quickly ensues and the Picture Theorist is stuck with an infinite chain of images, each one taking as its intentional object another image.

As I see it, there is a plausible reply to the second argument, although I see no plausible reply to the first. The reply to the second argument is that sense-data are the basic objects through which we experience the world. There is no reason the sense-data theorist has to commit to the principle that our awareness to sense-data is only by proxy. He need claim only that our awareness of sense-data is primitive, which is not only an assumption sense-data theorists make, but it is also the primary motivation for the sense-data theorist postulating sense-data: so that there is a type of evidence which is basic and incorrigible.

In any case, I think the Medium Objection and Sartre’s work are enough to show that the illusion of immanence is to be avoided. McGinn prescribes a naïve direct realist view, which he summarizes as:

In the case of images I would say, boringly enough, that they are experiences in the having of which we apprehend external objects—though, of course, the mode of consciousness involved is imagistic, not perceptual. (McGinn 2004: 68)

Materials Used in Imagination and Conceiving

There seems to be a rather obvious objection to this view. It looks much too naïve when it comes to imagining complex situations, which we do in thought experiments. Someone given to possible-worlds talk might reject the theory on the following grounds:

5. We can imagine possible worlds.
6. Possible worlds are not external objects.
7. Therefore, what we imagine are not just external objects.
8. Therefore, the naïve theory is false.

(5)’s truth will depend on how we understand ‘possible worlds’. If all we mean by this locution is something like ways the world could be, then I do think that we can, and many times
do, imagine possible worlds. But if we mean something more ambitious, then I am skeptical of our ability to imagine such worlds. Understood in the most straightforward way, (6) is true. Whatever possible worlds are, they should not be confused with mere actual objects. I grant this, but on my view possible worlds, to fully lay my cards on the table, are mental representations of the actual world. When I ask you to imagine a world like ours in every way except it has no France, your ability to imagine this word depends on thinking about actual objects, properties and arrangements of them. It is not as if you are grasping an abstract object when you imagine a France-less world or a Lewisian possible-to-us-but-actual-relative-to-itself world (Lewis 1986).

The naïve theory seems like a candidate for explaining how it is we can imagine the many diverse things we do, as opposed to the view that possible worlds are abstract objects we grasp or concrete entities outside of our realm of space and time but in a realm of space and time of their own (Lewis 1986). I also think that a representational account can be put to work here, but we also do not want to be committed to the view that we are direct realist about imagination, for we can imagine things that are not actual objects. For example, I can imagine a man with eight heads. The heads go straight up. I have a certain representation that is not directly about an object in the sense that it directly pictures it, but it is directly about it in the sense that it is of a particular person, standing in a particular spot, and particular heads from different folks. Each bit is about something actual, even if the whole is not.¹⁶ We should leave enough slack in the direct realist account to have that sort of story. We do not have to think of it as on par with direct realism with regard to perception.

¹⁶ Wittgenstein seems to express some sympathy for this view. In The Blue and Brown Books he writes,

“How can one imagine what does not exist?” The answer seems to be “If we do, we imagine non-existent combinations of existing elements”. A centaur doesn’t exist,
The Happy Consumer Fallacy

This fallacy centers on the fact that a subject’s descriptions can misdescribe what the actual representation is about. The vivid nature of mental images and the ease with which we generate them gives us a false confidence as to their accuracy. Imagine, for example, that the moon is made of green cheese. You may very well have a coherent image of a green moon. You seem to get the image with much ease. There seems to be nothing that would rule out such a possibility, at least based on your image. But clearly there are other factors that need attending to. As van Inwagen notes, there are crucial details missing from your representation of the target situation. But you think it is coherent nonetheless. This I chalk up to not necessarily the illusion of immanence but something like it, say, the happy consumer fallacy: Someone is so happy about a great deal that they forget (or do not want) to read the fine-print.

I should slow down and reiterate an above point. There is a sense in which people do have first-person authority over the product of their own imaginative processes. They have the final say, trivially, on their descriptions of the target situation. But surely they are not the final arbiters of whether that product is coherent or even accurate. But the link between target and subject’s description is so tight that there is a huge presumption in favor of accuracy. The tight link leads us to neglect to read the fine print.

First-Person and Third-Person Perspective in Imagination

Initially I said that first-person acts are acts which involve a subject essentially. Here is a more formal definition:

---

but a man’s head and torso and arms and horse’s do exist. “But can’t we imagine an object utterly different from any one which exists?”—We should be inclined to answer: “No; the elements, individuals, must exist. If redness, roundness and sweetness did not exist, we could not imagine them.” (Wittgenstein 1960: 31)
An act of imagination is first-person act iff there is essentially a subject’s perspective in the product of the imaginative process.

Let us see how this first pass measures up against distinctions drawn in the literature. In “What it is like to be a bat” Nagel offers the following:

To imagine something sympathetically, we put ourselves in a conscious state resembling the thing itself. (This method can be used only to imagine mental events and states—our own or another’s.) \(^{17}\) (Block, Flanagan and Güzeldere: 527)

So, for Nagel, to imagine from the first-person point of view, or sympathetically, is to put oneself in the mental state that is like the thing we set out to imagine. The most obvious observation here is that we can imagine things from a first-person point of view only if those things have conscious mental states. I cannot imagine a rock sympathetically. While it is another question whether or not I can imagine a bat sympathetically, suffice it to say that we can sympathetically imagine only situations in which there is a mental state in the product.

Sydney Shoemaker says that the imagining must be “done from the inside” (1993). And this seems by and large correct. Intuitively, first-person acts are those in which one is at the center of the situation. One is the subject that is experiencing the situation. When we imagine from the first person perspective we imagine experiencing something. We instantiate phenomenal properties in virtue of the perspective we take in the imagined situation (the product). Trivially, we will always instantiate phenomenal properties when we imagine something, but imagination is done from the first-person perspective when those properties are in the product essentially, as I tried to suggest in Chapter 1.

\(^{17}\) Although it is not clear from the quote, Nagel’s definition does not rule out that we can imagine mental features, like the property of being conscious, from a third-person point of view, which is obviously of interest to my project and bears on Marcus’s worry. I certainly seem to be able to imagine that someone else is conscious even if I do not put myself in that someone’s shoes. In fact, this is something we do all the time. Every time one reads a novel they imagine characters being conscious from a third-person point of view. I may imagine what it is like to be them, but I certainly need not in order to think of them as conscious.
One also need not think that the subject in the imagined situation essentially need be identifiable with the imaginer. We need not imagine that we are the person in the product of our imaginative process. To sympathetically imagine George Bush is to imagine being George Bush from a subjective point of view that somehow resembles his own—full stop. I do not have to imagine being the man himself.\(^{18}\)

I think this distinction is intuitive enough. However, some think that all imagination is done from the first-person perspective. Peacocke, like Berkeley before him, thinks this is the case. Allow me a brief response to this view, since it does stand in the way of the intuitive distinction I wish to draw. Peacocke thinks that all products of imagination involve a subject essentially:

\[
\text{imaginings always involve imagining from the inside a certain \textit{[type of] viewpoint, and someone with that viewpoint could, in the imagined world, knowledgeably judge \textit{“I’m thus-and-so,”} where the thus-and-so gives details of the viewpoint. (Peacocke 1985: 20)}
\]

This is a mistaken view for a couple of related reasons. First, it confuses the process of imagination with the product of imagination.\(^{19}\) Peacocke’s reasons for thinking this is based on his notion of S-imagining mentioned earlier. Just to review a bit, Peacocke develops this notion of S-imagining to spell out what it is that distinguishes different products of imagination that share a common image. S-imagining is like supposition; deploying it adds background conditions to the representation that are not explicitly contained in the image. Peacocke gives the example of imagining a suitcase and, then, imagining a suitcase with a cat

---

\(^{18}\) Shoemaker makes this point (1993).

\(^{19}\) Bernard Williams in “Imagination and the Self” makes this general line of criticism against Berkeley that I think is the right thing to say to Peacocke as well. I do not have the space to develop it here, though (Strawson 1969: 192-213).
inside. What makes these two images different representations is that in the latter case we S-imagine that there is a cat in the suitcase.\textsuperscript{20}

Now, Peacocke thinks that the S-imagined conditions we need to apply in imagining a situation involve supposing that someone is experiencing the situation. This supposition rules out the possibility of imagining a tree unperceived because the perceiver needs to be supposed to be in the imagined situation in order for the product to correctly represent a tree being in the quad. Here is what Peacocke says,

In being asked to imagine an unperceived tree, we are asked not just to imagine the sort of experience one has when one sees a tree, but to imagine a tree, really there in front of us. What this last involves, I have argued, is that the imaginer not merely imagine from the inside an experience as of a tree, but also that he S-imagines as a condition on the same imagined world that the experience is a perception of a tree. So when he imagines a tree, the S-imagined conditions entail that, in the imagined world, some tree is perceived. To combine this with the supposition that in the imagined world no tree is perceived is to place inconsistent conditions on the imagined world: and that was precisely Berkeley’s conclusion. (Peacocke 1985: 28)

So, the S-imagined conditions entail that there is a subject essentially in the imagined situation.

In reply, I think it is not an entailment we should accept, because we do not perceive the tree at all. In fact, we seem free to suppose that no-one is in the imagined situation using Peacocke’s own notion of S-imagining. If there is no contradiction in supposing that there is no-one in that situation, then one properly imagines a tree unperceived. It seems too much to demand that we must stipulate that there is a perceiver in the situation to ensure that there is a tree there. That, frankly, seems necessary only if one assumes that idealism is true. I hope that if one were tempted to think that all imaginings are done from the first-person, one no longer is based on the reply to Peacocke’s case for that very claim. Further, I hope to have

\textsuperscript{20} This is not the only way to imagine a cat in a suitcase.
moved the first-person and third-person somewhat past the mere intuitive distinction in this section.

**Chalmers’s Distinctions**

It is necessary to address three distinctions Chalmers makes with regard to conceivability (2002). The first is between negative and positive conceivability which marks acts of imagination with some sort of “mediated objectual character” off from those that lack this character. The second, between prima facie and ideal, concerns whether or not the act of conceivability in question can be defeated by better reasoning. The last distinction, between primary and secondary conceivability, is made about the mental representation we have in imagining itself and made when the imaginer is deciding whether the imagination is coherent. In a bit more detail, to determine primary conceivability we ask: might this imagined world be the actual world? If we answer yes, then it is primary conceivable. If no, it is not. As for secondary conceivability, one asks, can I coherently describe this situation as one that might have been the case? If so, yes. If not, no.

**Positive vs. Negative**

The first distinction I discuss is between positive and negative acts of conceiving. This distinction is related to what I think is the dominant conception of the distinction between imagining and conceiving. The idea here is that imagining is like perceiving and conceiving is not. This present distinction does not track the dominant conception entirely, but we can draw some parallels.

Positive conceivings are those in which “one can form a positive conception of a situation in which S is the case” (Chalmers 2002: 150). The representations involved are close to perceptual imagination, where we have a representation in mind that, in Chalmers’s words, usually has a “mediated objectual character”. In this essay I adopt this terminology.
Of course this is not to say that the representation need represent every detail of the target situation. For example, if I ask you to imagine a tiger with 211 stripes, your representation will have a certain objectual character, but it will not represent each stripe.

To see how this is to be distinguished with negative conceivability, let us think back to our mile-high unicycle. If you have a representation of what such a world would be like—no matter how rough—then it is a case of positive conceiving. But if you proceed in a way such that you reflect on the concepts of unicycle and mile-high and do not have a representation with any sort of mediated objectual character, then you are negatively conceiving. If we do not actually “picture” a ridiculously huge unicycle, then we negatively conceive. Positive conceiving is closer to the traditional conception of conceiving, which is to have a Cartesian clear and distinct idea of something. Negative conceivability, however, does not involve this modal “seeing”. Situations that are negatively conceivable, in Chalmers's words, “can not be ruled out as false a priori” (149).

We might say then that when one negatively conceives he is engaging in a thin sort of conceiving. This sort of conceiving involves not much more than logical reasoning about the sentence representing the situation in question. At a certain stage in imagining Kirk’s zombie we had to go negative: we could not get a positive picture of the absence of phenomenal properties in Dan’s head, we just had to live with the fact that we could not rule out the possibility that he is unconscious.

---

21I think there is a skeptical case to be made against using negative conceivability for anything more than the simplest sort of situations. I think it appeals to an unrealistic view of how we can examine our concepts. But I pass over this point here.

22This consideration suggests that there are mixed cases: products that are arrived at by positive and negative processes of conceiving.
In “Conceivability and the Cartesian Argument for Dualism” James van Cleve (1983) discusses two types of conceivability (Strong and Weak), which inspire Chalmers’s distinction between positive and negative conceivability. To explain this distinction, we need to discuss the technical sense of “see” that van Cleve introduces. Philosophers, he says, often say that they just “see” that certain propositions are true, simple mathematical truths for example. When we see that 2+3=5, we, in Descartes’s words, have a “clear and distinct perception” that this is so (Van Cleve 1983: 36). Using this notion of “seeing,” van Cleve gives his definitions of strong (like positive) conceivability and weak (like negative) conceivability:

- **Strong**
  \[ P \text{ is strongly conceivable for } S \text{ iff } S \text{ sees that } P \text{ is possible.} \]

- **Weak**
  \[ \text{If } P \text{ is a proposition that } S \text{ is considering, then } P \text{ is weakly conceivable for } S \text{ iff } S \text{ does not see that } P \text{ is impossible.} \]

Van Cleve talks of the conceivability of propositions. This may sound at odds with my view that we conceive of situations. I do not think there is anything to worry over here. Most propositions are only conceivable or inconceivable in so far as we have an idea of what situations they represent. Often one may say that a certain proposition is conceivable. But to say that is just to say that the situation the proposition represents is conceivable.

**Is Conceivability a Better Guide to Metaphysical Possibility than Imagination?**

In his *Meditations*, Descartes says that we can conceive, but not imagine, a chiliagon. The point here is that we can evaluate the status of certain proposals without the benefit of the mediated objectual character we typically associate with imagining. Descartes clearly thinks that this sensory character hinders us from grasping certain metaphysical truths, e.g., *that matter is essentially extended*. In contrast, I think that imagining and conceiving are the same
fundamental process. My view is at odds with the Rationalist claim that rational faculties like understanding and conceiving are reliable guides to true beliefs, while perception and imagination are ultimately unreliable. As far as I can see there is no good reason to believe this. We find ourselves in errors of perception and errors of reasoning all the time. Just as one man has better vision than another, so one man has better reasoning faculties than another. Indeed, we find ourselves in errors of reasoning no less than errors of perception.

As I see it, there is nothing mysterious going on in conceivings that is lacking in imagination. As far as I can see they are not two different ways in which we represent the world. Conceiving does not tap into a faculty from which imagination is blocked. Imaginations merely have mediated objectual character—they are like muted perceptions—while what we might call “pure conceivings” do not. I acknowledge that in certain cases there need be something more than the pure image (cf. Descartes’s chiliagon, the world where numbers do not exist), but we need not think these are superior to what we might call pure imaginings. For example, the work traditionally thought to be done by conceiving is done by supposing certain conditions to hold in the imagined situation—the mechanism of supposition—and this mechanism when added to a theory of imagination captures any phenomenon we thought of as “pure conceiving.”

In any case, the mediated objectual character of imagination is what lead certain Rationalists to associate imagination with sense perception, and based on observations of the unreliability of the senses, to cast doubt on the reliability of imagination. I suggest we reject this distinction, if not on the reasoning above, then because we do not have a good explanation of what this supposed rational faculty comes to. I think a complete theory of imagination would cash out all the differences by pointing to differences at the level of the
concepts and other “devices” that get deployed in these putatively distinct ways of modal representation. In sum: there is no need, at least not one I can see, to posit another faculty.

Is Conceivability Equivalent to Negative Conceivability and Imagination to Positive Conceivability?

Now certainly there is a distinction between the phenomenologies of negative conceiving and positive conceiving that is similar to the distinction that many make between conceiving and imagining. I do not, however, think it would be right to say that “conceivability” is always of the negative sort. This, for starters, would do disservice to Descartes’s connection between conceivability and clear and distinct ideas. When we negatively conceive we do not have a clear and distinct idea. This, I believe, has led Yablo to reject negative conceivability as a genuine type of conceivability (1993). Not much hangs on this, but we do need to note that negative conceivability (or weak conceivability in van Cleve’s terminology) certainly is a grade below positive conceivability in terms of its link to possibility. The idea here is that having an idea of how a certain situation would be is a better guide to possibility than just being unable to rule out that situation a priori. This I take it is something on which most would agree.

Further, we need to be clear that cases where there is no definite mediated objectual character are not always negative conceivings. There seem to be two sorts of positive conceivings, the ones that have mediated objectual character and those that do not. The latter is exemplified by Descartes’s demonstration that extension is the essence of matter. When we “see” this, there is a certain point where the sensory character just runs out, yet we have some positive conception of that which we are to imagine. The former is exemplified by many sorts of imaginings.
Negative Conceivability and Skepticism about Thought Experiments

There is a connection between negative conceivability and skepticism about thought experiments. Some skeptical responses to thought experiments are driven by the idea that thought experiments are less like intuitions and more like implicit arguments. Acts of negative conceivability look very similar to straightforward arguments. In these cases we reason about situations without having a solid representation of what the given situation is like. Again, a proposition is negatively conceivable if we lack a positive representation of what that situation would be like.

Let us consider the example of imagining what it is like to be a bat. Ask yourself whether it is conceivable that there is something it is like to be a bat. It seems that the only way one can succeed will be by way of negative conceiving: we simply cannot have a positive conception of the experience of echolocation. But we cannot rule out such a situation a priori: for all we know, there is no privileged connection between our particular sensory modality and experience.

Daniel Dennett thinks that thought experiments proceed by way of negative conceiving and, thus, are more like implicit arguments and less like clear and distinct ideas. As such, he thinks that thought experiments do not reveal intuitions but just theoretical assumptions—like an argument once it is laid plain.

Let us look at how Dennett thinks this works. He argues that in order to imagine zombies one must presuppose that there is a distinction between p-consciousness and a-consciousness (Dennett 1999). If one presupposes there is a distinction, then Dennett thinks
it follows that zombies are conceivable. But Dennett thinks to make that presupposition is to beg the question:

To see the fallacy, consider the parallel question about what the adaptive advantage of health is. Consider “health inessentialism:” for any bodily activity b, performed in any domain d, even if we need to be healthy to engage in it (e.g. Pole vaulting, swimming the English Channel, climbing Mount Everest), it could in principle be engaged in by something that wasn’t healthy at all. So what is health for? Such a mystery! But the mystery would arise only for someone who made the mistake of supposing that health was some additional thing that could be added or subtracted to the proper workings of all the parts. In the case of health we are not apt to make such a simple mistake, but there is a tradition of supposing just this in the case of consciousness. Supposing that by an act of stipulative imagination you can remove consciousness while leaving all cognitive systems intact—a quite standard but entirely bogus feat of imagination—is like supposing that by an act of stipulative imagination, you could remove health while leaving all bodily functions and powers intact. If you think you can imagine this, it’s only because you are confusedly imagining some health-molecule that might or might not be present in a body. Health isn’t that sort of thing, and neither is consciousness. (Dennett 1999: 325)

Those who imagine zombies are already assuming that there is no such relationship: that mental states are not causally responsible for behavior. Dennett clearly thinks that this process is not a genuine guide to modal knowledge.

One may doubt that the question can be begged simply by imagining or conceiving of a situation, since, strictly speaking, the things that beg the question are arguments that contain premises (implicit or explicit) that are the same as the conclusion. But to see the force of the objection, let us go back to Dennett’s example. It provides a model for what is negative conceivability of the question-begging sort.

1. Health inessentialism is true. [Question-begging assumption]
2. It is conceivable that an unhealthy person could function as a healthy person. [from 1]
3. Healthy function is not essential to performing in a healthy way. [from 2]
4. Therefore, there is something to acting healthy that is not captured by healthy functioning. [from 3]

I think someone could presuppose there was a distinction between a-consciousness and p-consciousness and still find zombies inconceivable.
This argument begs the question, and it is hard to see how its conclusion could be true, since there does not seem to be anything more to being healthy than functioning healthy. If something like this is going on in the zombie case, we should not use the conceivability of zombies in support of any philosophical conclusion.

But is the zombie argument analogous to the health inessentialism case? Making the relevant changes, Dennett is thinking that those who want to argue from the conceivability of zombies to the falsity of physicalism argue thus:

5. Functionalism is false about phenomenal states.24
6. It is conceivable that a functional duplicate of us at the time we are in pain, might lack pain. [from 5]
7. Sameness functional state does not guarantee sameness of functional role. [from 6]
8. Therefore, there is something to conscious experience that is not captured by functionalism. [from 7]

Same bad argument, same reason. If the zombie advocate negatively conceives of zombies in this way by appealing to these sorts of premises, then Dennett can claim to have successfully deflated the intuition by showing that it is the product of theoretical bias.

I think Dennett is wrong on this point, though, because one need not run the thought experiment in the way he outlines. One way to see that the zombie advocate can negatively conceive of (8) being true is to first note that he need not help himself to the problematic assumption (5). To reason a priori from (6) and (7) to (8) is sufficient to show that—a priori—functionalism is false about phenomenal consciousness—provided we understand the meaning of the relevant terms. So, on a bit of reflection, it is easy to see that it is wrong for Dennett to think that (5) drives the thought experiment. It is sure to happen in some cases, but the fact that it need not and often does not is enough to show that this objection misses the mark.
What is more, Dennett’s way of thinking of thought experiments leaves us with an odd view when it comes to people who find that a situation is coherent yet not possible. On his view, such a person is using an implicit argument (thought experiment) to get to an implicit conclusion (intuition) but is also using an explicit argument to arrive at the negation of the above implicit conclusion. The Type-B Materialist is a good example: he thinks the zombie idea is coherent, but that zombies are impossible. Dennett cannot account for this response on his view. If the Type-B view is coherent, it looks like the reasonable thing to say is that intuitions are one thing, arguments another.

This is not to say that assumptions cannot run interference in conceiving. Suppose I tell an acquaintance that my friend Jesse (whom the acquaintance does not know) gave birth to a baby girl. Acquaintance says, “Surely, what you say is impossible, inconceivable even. How could a man give birth to a girl?” Well, the reason Acquaintance thinks that this situation is inconceivable is that Acquaintance assumes that Jesse is a man. Once Acquaintance rids himself of this false assumption, then there is surely no bar to conceiving of the possibility. The true intuition is that such a situation is conceivable, even if Acquaintance does not have a clear idea of who Jesse is.

This shows that beliefs (since assumptions are beliefs) have a role to play in the conceivability of situations. In addition to whatever image Acquaintance has when he tries to imagine a man giving birth is the belief that Jesse is a man. In fact this belief is prior to the imaginative act. Beliefs about what we are to imagine surely help structure imagination. Many times these beliefs turn out to be fine, but sometimes they are assumptions that run interference, as in the healthy-inessentialism and Acquaintance case.

So I grant that Dennett is half right: Implicit assumptions can cause us to have certain intuitions. But if we are careful we can show the difference between a genuine intuition and
a question-begging assumption. The health inessentialism case shows this. If one assumes that the property of being healthy is not necessary to functioning in a healthy manner, then one will find scenarios of unhealthy people performing in a healthy way conceivable.

**Prima Facie vs. Ideal**

The distinction between prima facie and ideal conceivability hinges on the imaginer’s cognitive ability. It is best to think of this in terms of an agent’s attention to a representation, even though it is a potentially confusing way of putting things. I do not mean to suggest a picture of imagination that I have already argued against: the picture theory. It is not as if an agent ranked high on the attention scale if he carefully examined his representation as if it were a museum painting. That is a false picture. What I mean here is the amount of effort the agent expends in examining the imagined situation. Chalmers writes that a situation is prima facie conceivable if, and only if, “after some consideration, the subject finds that S passes the tests that are criterial for conceivability” (Chalmers 2002: 147). Prima facie conceivable situations can turn out to be inconceivable in light of neglected information or can turn out to be genuine conceivabilities. The standards, though, for this act are not that demanding.

On the other hand, to ideally conceive is difficult, and requires sustained attention over a period of time. I do not mean that one must remain in the same act of imagination. Rather, one must turn on occasion to the same situation and “turn it over” in one’s mind, in order for it to be ideally conceivable. Ideal conceivability is done by a rational agent whose justification, as Chalmers writes, “cannot be rationally defeated” (Chalmers 2002: 148). Chalmers is careful to note that we are not to build any link between ideal conceivability and possibility into ideal conceivability such that it would trivialize ideal conceivability. The
conditions for determining ideal conceivability are independent of possibility. In the rest of
the essay I focus solely on ideal conceivability.

**Primary vs. Secondary**

Chalmers offers the following definition of primary and secondary conceivability:

We can say that S is primarily conceivable (or epistemically conceivable) when it is
conceivable that S is actually the case. We can say that S is secondarily conceivable (or
subjunctively conceivable) when S conceivably might have been the case. (2002: 157)

This category allows us to make a distinction between situations that are conceivable
prior to reference-fixing and inconceivable after these facts are fixed. For instance, on the
familiar Kripke-Putnam line about a posteriori necessities, water not being identical with
H₂O is primarily conceivable but not secondarily conceivable, because the reference of
‘water’ to H₂O is fixed a posteriori. So the idea is, for all I know a priori water might not be
H₂O, in the sense that I cannot rule out the possibility of that water has a rogue micro-
structure when I conceive in ignorance of its actual microstructure. But, in terms of a
posteriori conceiving, I conceive in view of the fact that “water” picks out H₂O necessarily.
So, no matter what sort of story I tell about water turning out to be a different substance,
*that* new substance will not be water in virtue of the claims of rigid designation and reference
fixing.

Because Chalmers has a special sense of “actual” in mind, it is not entirely clear that
his description of primary conceivability as what *might actually be* the case is the most helpful
way to characterize the phenomenon he is trying to capture. What we ask to determine
primary conceivability is “Might this actually be the case?” but we do not have any world in
mind that ‘actually’ indexes. Yablo (2002) suggests using “counter-actual” since this seems to
capture that we are not concerned with our world but a possible world we are considering as
if it were the actual world.
The following comments, from Chalmers, shows that Yablo and Chalmers disagree about whether Chalmers’s notion of primary conceivability is conceivability:

(Primary conceivability is related to what Yablo (1993) calls “conceivability ep,” which requires that one can imagine believing something true with one’s actual P-thought, but it is not quite the same. One difference is that primary conceivability does not require that a conceived situation contain a P-thought. So it is primarily conceivable that nothing exists, or that no-one thinks — these are not ruled out a priori, and are verified by certain situations considered as actual — but they are not conceivable in Yablo’s sense.). (Chalmers 2002: 158)

How is it conceivable that a world where nothing exists might actually be the case, unless we understood “actual world” in a way that rules out that our world is the actual world? Surely it is inconceivable that nothing exists in our world: merely reflecting on the question rules it out. Chalmers wants primary conceivability to be a purely a priori enterprise, but I want to suggest, in Chapter 4, that this is a problem.

Since this is an issue that comes up later, let us pause to see if we can get more help. Clearly we cannot merely ask whether it might actually be the case, because that would rule out the conceivability of a situation verified by the proposition that no-one thinks. Chalmers writes,

When we consider situations as actual, we consider and evaluate them in the way that we consider and evaluate epistemic possibilities. That is, we say to ourselves: what if the actual world really is that way? One hypothetically assumes that the situation in question is actual and considers, whether from that assumption, it follows that S is the case. If so, then the situation verifies S, when considered as actual. (Chalmers 2002: 158)

So, primary conceivability consists of judging whether a certain situation is coherent, and that situation need not be our world. Instead of asking “What if the actual world really is this way?” we need to ask, on Chalmers’s definition, “What if a world is this way?” So, primary conceivability depends not on actual facts. It merely concerns assuming that a given situation is actual, and then asking if the situation which one set out to conceive of might be the case in that world.
How might these above concerns carry over to the case of thinking about phenomenal properties and physical properties? If water not being H₂O is secondary inconceivable, then why not pain’s not being C-fibres firing? First, we should note that a situation in which all the physical properties of the actual world are instantiated but no phenomenal properties are will be primarily conceivable even for the physicalist, because the reference-fixing claim of the identity theorist that “pain” refers to C-fibres firing is not something we can attend to. It seems primarily conceivable—so the story goes—that there may be all those physical properties and no phenomenal ones, as we saw in the water case. But the physicalist can mimic the water-monist move in adopting Kripke’s story about “water” referring to H₂O in all possible worlds. So, the physicalist makes the analogous move in the pain case: pain without C-fibres is secondarily inconceivable because “pain” is a rigid designator that refers to C-fibres firing in all worlds. However, according to Chalmers, this will not work because we do not think that the secondary extension of “pain” is a physical property like being H₂O. The intuition that Chalmers and Kripke want to press is that the extension of ‘pain’ is the feeling of pain and not a property of the brain. Unless we can show that intuition is mistaken, saying that zombies are secondarily inconceivable seems like a dead end.

**Summary**

In this section we discovered that the structure of imagination and conceiving is such that there is a presumption in favor of our mental representations accurately representing the target situation due to the tight connection between target situations and subject’s descriptions. We also saw that imagination is deceptive in ways that lead us into modal error (the illusion of immanence and the happy consumer fallacy). We discussed perspective in imagination and Chalmers’s distinctions (2002). I raised a worry about primary conceivability
that I hope to show (in the next Chapter) changes the zombie thought experiment rather
dramatically.
CHAPTER 4
PULLING THINGS TOGETHER

Some of the phenomena we canvassed in Chapter 3—the illusion of immanence, happy consumer fallacy, and seeming infallibility of subjects’ descriptions of the products of their imaginative processes—should give us some pause with regard to our imaginative abilities. It does not look to be the case that we have ultimate first-person authority over whether the product of our imaginative process correctly represents what we set out to imagine, since we have seen that we can be mistaken. Of course, we do have undeniable first-person authority about how the products of our imaginative processes seem to us. But we can be mistaken about what they actually represent. At times, as I have tried to show, we are blind to some of the deeper aspects of the products of our imaginative processes and cognitive mechanisms at work. In short, we are sometimes mistaken about the success of our imagination. Often it is the phenomenology and ease of imagination that runs interference. I hope to have made a case for taking imagination more seriously: to engage in what we might call, following van Inwagen, “the active imagination.”

Two relevant examples discussed were the imitation man and the super-blindsighter. In the case of the imitation man, we saw that when we start looking at the details of the situation, the imitation man does instantiate phenomenal properties in virtue of his awareness of his pains. This follows from the plausible assumption I laid out at the end of Chapter 1. The imitation man does not, to be sure, share our normal phenomenology, but he is p-conscious in virtue of awareness. To deny this seems to commit one to a false view of
what p-consciousness is. Many may be tempted upon first imagining the imitation man to follow Campbell’s suggestion that the imitation man is not p-conscious, but this is a mistake.

One way in which we imagine the super-blindsighter presents another case in which we are duped into believing we have succeeded in imagining something when we have not. For instance, we saw that one must imagine the mechanism by which the super-blindsighter receives verification that his a-conscious states correctly depict the world. If we do not imagine this mechanism, or suppose he has such verification, the super-blindsighter would not be able to use that content in a rational way. He would not have any reason to use it in a premise in reasoning, because his a-conscious states would be, for all he knows, unreliable guides to how the world is.

We learned further lessons from the super-blindsighter. Actively imagining the super-blindsight case shows us that the super-blindsighter is not a complete zombie, because he is aware of his environment in a way that makes him p-conscious—accepting, of course, the plausible assumption from the end of Chapter 1. Block accepts this assumption, but, curiously, thinks that zombies are a-conscious but not p-conscious. This seems to be a mistake. For how are we to think of creatures that are aware of their beliefs and not p-conscious in virtue of this awareness? I can think of no way to do this. I should note, though, that this is not devastating for the Zombie Argument. Imaging, say, color zombies (creatures who behave just as we do and are physical identical to us but do not see colors) is still sufficient to challenge physicalism, as Chalmers notes. But many times zombies are described as experiencing nothing. These latter sorts of cases, I think, we have good reason to rule out, for the reasons just given.¹

¹ Out of historical interest, I should mention that Saul Kripke and Hilary Putnam developed a way to show that imagination is not always a guide to possibility. Imagining that water is
Again, I think these points suggest we should take imagination more seriously. A tentative conclusion I would like to draw based on these observations from the foregoing sections is that imagination is not incorrigible in the way we are tempted to believe, mainly because we are sometimes blocked from seeing all the details of the situation we are targeting. Below I try to make good on this one more time by showing how Chalmers’s notion of primary conceivability demands that we change the zombie scenario in a rather surprising way. Primarily conceiving of zombies turns out to be more demanding than we may have previously thought.

Before getting to this final worry, let us pick up another dominant thread running throughout the essay, which is the issue of imagining zombies from a third-person point of view and the question of whether this was a genuine way to imagine zombies. In short, I think that Marcus’s demand is too demanding. I think his objection, as well as Hill’s worries about splicing, can be met by paying attention to the mechanism of supposition. A plausible rejoinder to this seems worth mentioning. Marcus and Hill can reply that supposition is itself an unreliable mechanism. However, to say this is to suggest that a good many of our non-controversial successful imaginations would be thrown into doubt because supposition (and XYZ presents an illuminating case study their method of misdescription. Putnam (1975) and Kripke (1980), who held that in the case of *a posteriori* necessities, the apparent contingency of identity statements using two rigid designators was shown to be merely apparent. For example, it might seem to us there could be water in a world without \( \text{H}_2\text{O} \), but what we were imagining in that case was not water but some other substance. The reason we cannot imagine such a world is that “water” is a rigid designator, a term which picks out the same referent in every possible world. Based on the actual world referent of “water,” we know that “water” picks out \( \text{H}_2\text{O} \) as a result of facts about its underlying microstructure, not as a result of superficial surface properties, or so the story goes. In the language Kripke uses in *Naming and Necessity*, being wet is an inessential property of water; being \( \text{H}_2\text{O} \) is an essential one, and the only essential one (Kripke 1980). What we imagine to be a water-full, \( \text{H}_2\text{O} \)-less world is actually a world with another water-like substance; we misdescribe it in using the term “water.” I relegate this to a footnote for two reasons. First, it has proven to lack application to cases of “pain” and “C-fibre firing.” Second, it rests on a semantic thesis to which not everyone agrees.
the related phenomenon of intending) play prominent roles in almost all cases of imagination. In tightening the standards on imagination, Hill and Marcus squeeze out the uncontroversial cases as well.

As I said, the project is not to defend any skeptical thesis about the conceivability-possibility link but to keep such a link in good standing because it seems to be our only non-trivial access to modality. At this point in the project, one may think that I have actually strengthened the zombie intuition. Nothing I have said seems to throw that intuition into doubt in any serious way. In the remainder of the essay I want to try to say why I think, based on Chalmers’s notion of primary conceivability, the zombie scenario looks very much like the other minds scenario briefly discussed in Chapter 1.

Chalmers’s master argument against physicalism relies on his notion of primary conceivability. So, if we can show that the primary conceivability of zombies is more difficult than we may have thought, we will press him on an important point (Chalmers 2002: 198). I shall argue that the primary conceivability of zombies comes down to the other minds scenario. If this worry is true, then the zombie scenario is not at all like most thought: as a situation in which no one is conscious. To develop this worry, let us start with a quote from Chalmers,

In the zombie argument itself, the claim is that it is conceivable that in the actual world \( P \) holds, but no one is conscious. (Of course, I know that I am conscious, but this is a posteriori knowledge; that issue can also be bypassed by considering only the epistemic possibility that \( P \) holds while others in the actual world are zombies. (Chalmers 2002: 1996)

Based on this, it looks like there is some tension in the claim that primary conceivability is an a priori notion, which is clearly an idea Chalmers is committed to. He notes that this notion is grounded in the idea that “for all we know a priori, there are many ways the world might be” (Chalmers 2002: 157). But from a purely a priori standpoint, we
seem to have a difficulty in locating which world is the actual world. Chalmers seems to
think that there are two ways of dealing with this tension. First, we could abstract away from the
a posteriori fact that one is conscious or, second, we could embrace it and then imagine that
everyone else is unconscious (the other minds scenario). I want to suggest that the first line
of defense does not work. I grant you that it seems conceivable that P holds and no one is
conscious. But that is only conceivable as the way a world might be, not the way the world is.
I must know that I am conscious in that world in order to identify it as the actual world. It
seems impossible that it might actually be the case that I am not conscious. After all, what
reason do I have to think a world where no one is conscious and P holds is the actual world
if I am not conscious in it? I can certainly conceive of it as a way a world might be, but not
the way this world might be, right now that is.\(^2\)

So, it seems that the closest one can get to a complete zombie scenario when one
primarily conceives is to suppose that one is the only creature conscious in the world one is
considering as the actual world. But if one is in the zombie world oneself that seems to
change things rather dramatically. In this sense we must imagine the zombie world, at least in
part, from a first-person perspective. In fact, not only do we need to face the familiar fact
that we lack positive evidence for zombiehood in others, but we also have to contend with

---

\(^2\) Note also that there is an interesting ambiguity in the way Chalmers sets things up. He says
that primary and secondary conceivability correspond, in his words:

to two different ways of thinking about hypothetical possibilities: epistemically, as
ways the world might actually be, and subjunctively, as counterfactual ways the world
might have been. (Chalmers 2002: 157)

‘The world’ is used differently in the two instances. In the first instance it refers to whatever
world we are considering. The world in question is not necessarily our world (although it
may be), but whatever world we are considering at the moment. So the possibility is relative
to that world. For possibility to be of interest to us it has to be across the space of all worlds:
we can only achieve that in the second sense where ‘the world’ refers to our world.
an argument against the possibility of only myself being conscious. This transcendental argument was given in different forms by Wittgenstein (1953) and Strawson (1959).

Here is where some of the comments about the other minds hypothesis become relevant. Consider the following transcendental argument against the other minds problem:

1. I know that I am minded. [Undeniable epistemic fact]
2. To know that I am minded, I must have the concept of being conscious, or consciousness we might say. [Necessary condition for knowing one is minded]
3. In order to have the concept of mindedness, one must be able to distinguish between one’s own states of consciousness state and others’ states of consciousness. [Claim about concept acquisition]
4. So, in order to have the concept of mindedness, there must be others who have minds. [From 3]
5. Therefore, asking the question whether or not others have minds presupposes that others have minds. That is, they must have minds for me to even ask the question. [From 2,3,4]

I do not reconstruct this argument because I think it shows that the primary conceiving of zombies is impossible in any obvious way. My goal here is to draw up the relevant “off-stage machina” we must contend with if we are to actively imagine zombies. Recall that to be primarily conceivable the situation must be conceived as actual—“as a way the world might actually be” in Chalmers’s words. Recall that in the zombie case, we must situate our consciousness in the zombie world if it is to be primarily conceivable. If we are to have a representation that we can plausibly take to verify the target situation, we need to have a way to do satisfy the follow sets of claims in a way that is consistent:

6. I am conscious in the zombie world.
7. In the zombie world my mental states are identical\(^3\) to some physical states in the brain.
8. My concept of self-consciousness is not parasitic on understanding what it is for others to be conscious.

\(^3\) Identity is not the only possible relation here. I suppose one might think that his mental states are reducible to or supervene on his physical states without any remainder. But when we think hard about these other sorts of relations they seem, at bottom, to be a relation of metaphysical identity, however one wishes to express it.
The reason (6) must be accepted, again, is that if we cannot locate our own consciousness in the world, then we cannot say that it is actual. To think that a world without conscious might be the actual world is to endorse a rather radical skepticism about our ability to know that we are conscious. This is not a skepticism in which we should trade.

The reason that (7) must be accepted is as follows. The thought here is from (6) and the stipulation that all the physical facts hold of our world hold, and the further stipulation that only the physical facts hold, which is needed to rule out the possibility that the zombie world contains non-physical stuff, we get an entailment to (7). In short, we know we are minded and we know there is just physical stuff due to the situation we are asked to imagine. Therefore, our minds are physical.

The reason we must be able to hold (8) is that we have to, if we are to actively imagine, rule out the worries raised by the transcendental argument mentioned above. But to accept (8) we must have an answer to the transcendental argument. We need a story that explains how we can understand the concept of mindedness in the absence of other minds. This is surely a condition placed on the primarily (and ideal) conceivability of zombies that is not clear before we started this investigation. In addition to this, we face an additional difficulty that is not present in the other minds case: we cannot tell a positive story about the other creatures’ internal constitutions that explains why they lack minds. In fact, we know that we are minded in virtue of physical states of the brain, and our fellow creatures have those same physical states. The zombie hypothesis seen in the light of the active imagination comes down to a stronger version of the other minds hypothesis.

I do not wish to try to enter the debate of whether the transcendental argument is sound. I just want to note that it is a real challenge to the primary conceivability of zombies if we take imagination seriously. I also want to emphasize that primarily conceiving of
zombies is more difficult than conceiving of a situation of which the other minds hypothesis
is true because I have no positive evidence that others are minded. It seems that I must
suppose that they lack minds. If I do this, they are primarily conceivable provided there is no
deep contradiction in my having the concept of self-consciousness in a community in which
no-one else is minded. At best this is in need of further explanation and at worst incoherent,
but a problem we must face up to when we primarily positively conceive nonetheless. At the
very least, this shows that primary positive conceivability is more difficult than Chalmers
makes out.
WORKS CITED


BIOGRAPHICAL SKETCH

Casey Woodling is completing his third year of graduate study at the University of Florida. He received a B.A. in both literature and philosophy from the University of Central Florida. His philosophical interests include philosophy of mind, philosophy of perception, and the role of imagination in philosophy.