Georgia State University
Digital Archive @ GSU

Philosophy Theses | Department of Philosophy

6-9-2006

Ways to Skin the Zombie Cat: A Look at the Problems Associated with Chalmers's Zombie-Argument

Walter Scott Clifton

Follow this and additional works at: http://digitalarchive.gsu.edu/philosophy_theses

Recommended Citation
Clifton, Walter Scott, "Ways to Skin the Zombie Cat: A Look at the Problems Associated with Chalmers's Zombie-Argument" (2006). *Philosophy Theses.* Paper 3.

This Thesis is brought to you for free and open access by the Department of Philosophy at Digital Archive @ GSU. It has been accepted for inclusion in Philosophy Theses by an authorized administrator of Digital Archive @ GSU. For more information, please contact digitalarchive@gsu.edu.

WAYS TO SKIN THE ZOMBIE CAT: A LOOK AT THE PROBLEMS ASSOCIATED WITH CHALMERS'S ZOMBIE-ARGUMENT

by

W. SCOTT CLIFTON

Under the direction of Reina Hayaki

ABSTRACT

In contemporary philosophy of mind, the issue of consciousness has taken center stage. Broadly speaking, those who deal with consciousness fall into two camps: those who prioritize empirical work and those who favor conceptual investigation. One prominent argument has served to deepen the divide: the argument for the possibility of zombies. In this paper I intend to examine closely this argument, as it's presented by David Chalmers, and some of the attempts to discredit it. In so doing, I present some of my own arguments against it, as well as the claim that if it's sound, then materialism is false. Finally, I present a sketch of a new way of thinking about consciousness that would, I argue, guard against the threat—real or merely apparent—of arguments such as the zombie-argument.

INDEX WORDS: Consciousness. Logical possibility of zombies. David Chalmers. Conceivability-arguments.

WAYS TO SKIN THE ZOMBIE CAT: A LOOK AT THE PROBLEMS ASSOCIATED WITH CHALMERS'S ZOMBIE-ARGUMENT

by

W. SCOTT CLIFTON

A Thesis Submitted in Partial Fulfillment of the Requirements for the Degree of

Master of Arts

in the College of Arts and Sciences

Georgia State University

2005

Copyright by
Walter Scott Clifton
2005

WAYS TO SKIN THE ZOMBIE CAT: A LOOK AT THE PROBLEMS ASSOCIATED WITH CHALMERS'S ZOMBIE-ARGUMENT

by

W. SCOTT CLIFTON

Major Professor: Reina Hayaki
Committee: George Rainbolt
Robert Almeder

Electronic Version Approved:

Office of Graduate Studies
College of Arts and Sciences
Georgia State University
December 2005

TABLE OF CONTENTS

INTRODUCTION…….1

THE ZOMBIE-ARGUMENT……5

THE POSITIVE ARGUMENT…….11

The Conceivability-Possibility Conditional……11

The Self-Undermining Feature of C-P……19

The Yablo-Chalmers Picture of Conceivability and Goldbach's Conjecture……23

Conceivability of Zombies……..32

THE BURDEN-OF-PROOF ARGUMENT……...37

THE INHERENT USELESSNESS OF THE ZOMBIE-ARGUMENT….44

A Deflationist Description of "Consciousness" and its Relation to the Zombie-Argument……...49

CONCLUSION……60

BIBLIOGRAPHY……63

Introduction

Daniel Dennett says that "consciousness is just about the last surviving mystery" (Dennett 1991, 21). David Chalmers echoes this sentiment: "Consciousness is the biggest mystery. It may be the largest outstanding obstacle in our quest for a scientific understanding of the universe" (Chalmers 1994, xi). Owen Flanagan, perhaps deliberately understatedly, writes, "The very idea of consciousness materializing, of subjectivity being realized in the activity of a physical organism, is puzzling" (Flanagan 1992, xi). And even Francis Crick, one of the pioneers in attempts to explain consciousness scientifically, admits that "consciousness appears so mysterious" (Crick 1994, 13). There are countless other statements which would highlight the odd nature of consciousness. It has been argued that anyone who wishes to take consciousness seriously has no choice but to admit the fact that consciousness is weird. To deny its weirdness is not to face up to the "hard" problem of consciousness.[1]

Consciousness is puzzling, mysterious, perhaps even an obstacle in our way toward understanding the universe. Until the last three decades or so, consciousness has been ignored. Toward the end of the nineteenth century, people like William James started to ask questions like "What is consciousness?" But these kinds of questions were by and large cast aside as irrelevant during the first half of the twentieth century. Attempts to explain the mental then primarily focused on behavior, or identity relations between the mental and physical, or causal relations between inputs and outputs. The what-it's-like aspect of the mental was often considered not a useful subject for study.

---

[1] (Chalmers 1994, xii)

But consciousness is now front-and-center in the philosophy of mind. Some embrace the mystery, opting to deny the possibility of our ever figuring out how consciousness fits in with our physical explanations.[2] Others have determined that the mystery *must* be illusory,[3] even if this fact has not been demonstrated; thus, it can safely be ignored. Scientific research will reign at the end of the day. It is up to us not to become entangled in the merely apparent mysteriousness of consciousness. Generally, opponents of those who hold this position object that these skeptics regarding consciousness are explaining consciousness *away*, rather than providing a scientific explanation of the phenomenal feel of experience.

There is a third kind of position to take—an intermediate position between defeatist mysterianism and skepticism. Of course, this "third position" cannot be described in a way that does justice to the very heterogeneous individual positions which are held. There are many theories which fall somewhere between the two poles. Perhaps one unifying characteristic of all these theories would be, however, a constant, unflagging consideration of consciousness and conscious experience. Whereas the mysterian position accepts the mysteriousness of consciousness at face-value and then gives up further exploration, and skepticism briefly fixes its attention on the phenomenal feel of experience and decides to ignore it for the sake of science, the intermediate position accepts that consciousness *must* be adequately explained, without being explained away, all the while admitting that there don't seem to be established ways to complete such an explanation. Thus, for this position, the mystery of consciousness is real, not merely apparent, but an explanation of consciousness is not, in principle, inaccessible to us.

---

[2] See (McGinn 1989) for a good example of this kind of position.
[3] Notice that Crick admits only that consciousness *appears* mysterious.

As we have already noted, the intermediate position is marked by diversity. There are those who believe that consciousness may be explained empirically, without much need for armchair conceptual analyses. Consciousness will be explained by science in some way—it won't need thought-experiments to guide it along its way. And there are those who still hope to explain consciousness primarily through conceptual analysis, i.e., by deriving truths about consciousness from closely inspecting the conceptual distinctions we make when we consider phenomenal experience.

One thought-experiment which has proved to be quite divisive among those holding the intermediate position is the argument for the possibility of zombies. While its history extends back thirty years, at least to Robert Kirk's seminal paper "Sentience and Behaviour," its popularity today may be attributed to David Chalmers and his book *The Conscious Mind*. Zombie-arguments have served a number of purposes—as an argument for epiphenomenalism,[4] an argument against functionalism, etc. Chalmers, however, uses his argument for the possibility of zombies as a way to refute materialism. If zombies are possible, he argues, then there is something in the actual world that is not physical—consciousness. Thus, not every object in the world is material. Therefore, materialism is false.

One can see here, then, just how significant the argument for zombies may be. If it succeeds, as Chalmers intends it to, then the larger argument against materialism succeeds. And if the argument against materialism succeeds, then Chalmers has shown that science, in its present state of trying to explain everything in physical terms, using strictly physical laws, will

[4] See (Perry 2001) for a discussion of Chalmers's zombie-argument as it relates to epiphenomenalism. See (Seager unpublished) for a reasonably good argument against a position such as Perry's.

fail to deal adequately with consciousness. Thus, science would have to be adapted to this purpose. Science would have to be overhauled, in order for it to be able to deal with consciousness.

It might seem to some utterly amazing that all of this can be demonstrated via something as abstract as a thought-experiment conducted from an armchair. Indeed, this lies at the heart of the divisiveness of the zombie-argument.

This paper has two purposes. The first will be to examine the zombie-argument and the place it occupies in Chalmers's argument against materialism. I intend to show that there are problems with the zombie-argument inherently, as well as the way it's used against materialism. The second purpose of this paper will be to present a deflationist description of consciousness, which, I will argue, allows for a more universal agreement about what we mean by the term "consciousness" and will allow more success in future investigations of conscious experience. I will further argue that utilizing the deflationist description of consciousness will also allow us to escape the clutch that the zombie-argument seems to have on us. As a full presentation of my deflationist view would require much more space than I have available in this paper, I will only be providing a sketch of my view, enough to make my point concerning zombie-arguments.

The final position I will take will obviously fall on the empirical side of the intermediate position. Science will eventually be able to explain consciousness, but it won't do so by explaining consciousness away. But until we get past having to respond to thought-experiments which try to take shortcuts to answers science should, in time, be able to provide, our time and effort will be wasted on speculation, rather than the hard work of research.

The Zombie-Argument

Chalmers presents his argument for the logical possibility of zombies in his chapter dealing with reductive explanation. He argues first that the logical possibility of zombies demonstrates that consciousness does not logically supervene on the physical. In a subsequent chapter, however, he refers to the zombie-argument in relation to his argument against materialism. His argument against the logical supervenience of consciousness upon the physical and his argument against materialism are closely related, but distinct. The goal of this section is to examine the zombie-argument in relation to the argument against materialism. Thus, we will have to draw from the zombie-argument as presented in the earlier chapter, but we will be considering it only in connection to the anti-materialist argument.

First, Chalmers provides the definition of "zombie" which he will be using. He defines a zombie—specifically, his own zombie-twin—thus: "someone or something physically identical to me … but lacking conscious experiences altogether. … 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). Thus, Chalmers's zombie-twin will be functionally, behaviorally, and psychologically identical to Chalmers. He just won't be having any phenomenal experience in cases where Chalmers himself would.

Second, Chalmers reasons from his being able to describe this zombie-twin to his being able to conceive of the zombie-twin. He claims that what he has described is coherent—there is no contradiction involved in what he has said. Moreover, his being able to conceive of this

zombie-twin implies that it is logically possible. He writes, “In some ways an assertion of this logical possibility comes down to a brute intuition … Almost everybody, it seems to me, is capable of conceiving of this possibility” (Chalmers 1996, 96). He then claims that some who deny the possibility do so only to make a theory come out right, when, in fact, the theory should conform to the possibility.

So, at this point it would seem that Chalmers’s argument for the logical possibility of zombies consists of two premises: (1) if zombies are conceivable, then zombies are logically possible; and (2) zombies are conceivable. The conclusion then is that zombies are logically possible. This I will call his positive argument for zombies. He has given us reasons to accept his premises; he does not expect us to accept his conclusion independently of his premises. Thus, he seems to believe that these premises establish his conclusion.

He then goes on to claim that there is really only one viable way for an opponent to object. The opponent would “claim that in describing the zombie world as a zombie world, we are misapplying the concepts, and that in fact there is a conceptual contradiction lurking in the description” (Chalmers 1996, 99). He does not think that objections regarding the reliability of conceivability arguments present many problems. He believes his first premise is pretty solid. He concedes that someone might object to premise 2 above, but he does not suspect anyone will present a convincing argument. He reiterates that he “can detect no internal incoherence; I have a clear picture of what I am conceiving when I conceive of a zombie” (Chalmers 1996, 99). Thus, he feels pretty good about the second premise, as well.

In addition to the positive argument for zombies, I would argue that Chalmers also puts forth a second argument for zombies. As we shall see, Chalmers needs the logical possibility of

zombies in order for his argument against materialism to succeed. He provides the positive argument above only as a means of establishing the logical possibility of zombies. There is also a sense, however, in which he argues that zombies are logically possible unless proven otherwise. He writes, "Arguing for a logical possibility is not entirely straightforward. … I confess that the logical possibility of zombies seems … obvious to me." He then goes on to say, "In general, a certain burden of proof lies on those who claim that a given description is logically *impossible*. … If no reasonable analysis of the terms in question points toward a contradiction, or even makes the existence of a contradiction plausible, then there is a natural assumption in favor of logical possibility" (Chalmers 1996, 96). Thus, he argues, zombies are logically possible until someone establishes that they are not.

Now, someone might object to my distinction of the positive and burden-of-proof arguments. It might be argued that the burden of proof he discusses is the same burden of proof on an opponent who objects to the second premise of what I am calling the positive argument. The contradiction which, he maintains, must be produced in order for zombies to be considered logically impossible would then be the same contradiction called for in connection with the second premise, the conceivability claim.

I would respond by saying that if what I am calling two arguments is indeed only one, then there does not seem to be any reason to point out that the burden of proof rests with Chalmers's opponent. If the entire argument for zombies consists of the two premises regarding conceivability and possibility, then it would be painfully obvious that the burden of proof rests on his opponent. But he clearly feels the need to point out the fact that the burden of proof is with his opponent. Also, why would he implement a strategy of, at one moment, providing two

premises in order to support his conclusion, and at another moment, claiming that his conclusion should be presumed true, unless proven false? It seems to me that these are two independent sets of support for his one conclusion. I would argue that this suggests that there are really two arguments, instead of only one.

Chalmers even suggests this in the text. After presenting his view that the burden of proof lies with his opponent, and before he discusses conceivability as a guide to possibility, he writes, "That being said, there are some positive things that proponents of logical possibility can do to bolster their case" (Chalmers 1996, 96).

There is very little attention paid to the burden-of-proof argument, perhaps because there is still so much to be said about the positive argument. Nonetheless, the burden-of-proof argument, we will see, is very important in relation to the argument against materialism.

So, we have seen how Chalmers presents the zombie-argument. Now all we have left to do is see how it operates in the anti-materialist argument. We can see that in the following—Chalmers's argument against materialism—the conclusion of the zombie-argument acts as a premise (Premise 2):

1. In our world, there are conscious experiences.
2. There is a logically possible world physically identical to ours, in which the positive facts about consciousness in our world do not hold.
3. Therefore, facts about consciousness are further facts about our world, over and above the physical facts.
4. So, materialism is false. (Chalmers 1996, 123)

His reasoning is that, if the zombie-world is physically identical to the actual world, but different in another respect (there is no consciousness in the zombie-world), then there is an extra fact about the actual world which cannot be a physical fact. “The character of our world is not exhausted by the character supplied by the physical facts; there is extra character due to the presence of consciousness” (Chalmers 1996, 123). Thus, not all of the facts about our world are physical. Therefore, materialism is false.

One thing we should note at this point is the strategy Chalmers employs in the above argument. It is an argument whose premises are strictly conceptual, except perhaps for premise 1. But premise 1 is, for the most part, uncontroversial. Premise 2 is a version of the conclusion of the zombie-argument, itself purely conceptual. Premise 3 depends on premise 2. And all three are needed in order to establish the empirical conclusion that there is at least one kind of thing in the world that is non-physical. But notice that with our assent to each premise, we are carried inexorably closer to the conclusion. If we assent to each and every premise, it seems we have no choice but to accept the conclusion. Indeed, we are forced into accepting the anti-materialist conclusion if we accept each premise. This is a characteristic of Chalmers’s argument we should note now, because it becomes relevant later.

Something else we should note now for future reference is who Chalmers’s intended audience is. It doesn’t seem to be dualists, at least not primarily. If it were, then he would certainly be expending a lot of unnecessary energy in order to construct an argument which forces its conclusion for an audience who already accepts the conclusion. The intended audience probably won’t be dogmatic materialists, either. The argument against materialism relies on the

zombie-argument and the latter relies heavily on particular intuitions which allow for the possibility of materialism's being false. Dogmatic materialists simply won't have the appropriate intuitions.

What about open-minded materialists? By this phrase, I mean those who believe in causal closure, scientific explanations, lack of overdetermination, etc., but who might still be unsatisfied with a purely physical reduction of consciousness. In a way, Chalmers himself seems to be an example of someone who has been convinced away from materialism, in spite of his belief in science. Much of his book focuses on the potentiality of discovering psychophysical laws which will allow for consciousness as a non-physical fact to be incorporated into science. Thus, it seems plausible that the open-minded materialist is an intended target.

And then there is, of course, the agnostic. Chalmers's argument is exactly the kind of thing someone who feels that there is not enough evidence to decide either way would be looking for. The fact that Chalmers's argument might convince such individuals as open-minded materialists and agnostics would certainly help to explain why his arguments have been hailed and closely examined in recent years. They have the potential to tip the scales of popularity back the other way, away from materialism and toward dualism, though his dualism is property dualism rather than substance dualism. Nonetheless, it's hard to deny that his work deserves the attention it has gotten.

Thus, Chalmers's zombie-argument has the potential to devastate materialism. For this reason, it hasn't been favorably received in all circles. Objections have been raised, especially regarding the positive argument. In the next section, I will mention a few of these objections, as well as raise a few of my own.

The Positive Argument

The positive argument has two premises, as we noted above. They are the conceivability-possibility conditional statement (referred to as C-P hereafter) and the conceivability claim. While Chalmers indicates that he does not think there should be much resistance to C-P, there are many who have, in fact, objected to this premise.[5] We will look at some of these objections first. Also, in spite of the fact that Chalmers insists that he cannot find a contradiction involved in conceiving of zombies, there have been many attempts to show that zombies are not, in fact, conceivable. We will look at some of these second.

The Conceivability-Possibility Conditional

Chalmers believes that the fact that he can conceive of zombies implies the further fact that zombies are logically possible. Shortly we will see exactly what he means by the terms "conceive" and "logically possible". Before we do, let's look at some attempts to call into question his prima facie claim that since zombies are conceivable, they are possible.

Christopher Hill and Brian McLaughlin (1999) accuse Chalmers of failing to provide even a rudimentary proof of the conditional. They claim that he merely responds to alleged counterexamples, but never provides a positive argument for its being true. Moreover, they

[5] "In our judgement, and probably in the judgement of most of the rest of Chalmers's readers, this principle [C-P] is the most vulnerable component of his view" (Hill & McLaughlin 1999, 449).

claim, there is good reason to think that when the Cartesian intuition (which they describe as "modal intuitions that *seem* to support property dualism" (Hill and McLaughlin 1999, 447)) is involved, the modal intuitions which seem to support C-P are not reliable. "As we have pointed out, there are significant differences between the cognitive factors responsible for Cartesian intuitions and those responsible for modal intuitions of a wide variety of kinds. Because of these differences, it would be irresponsible to generalize from other cases to the case of Cartesian intuitions" (Hill and McLaughlin 1999, 449). Now, it should be noted that Chalmers does provide a very detailed defense of C-P in (Chalmers 2002). Hill and McLaughlin's contention that Cartesian intuitions differ from modal intuitions is not addressed, however.

Another objection is raised by Dietrich & Gillies (2001). They argue, first, that Chalmers is committed to a Kripkean line regarding transworld identity; and second, that, given this commitment and Chalmers's notions of primary intension and secondary intension, he has no choice but to admit that zombies are not possible, even if they are conceivable.

First, they argue that Chalmers is committed to a Kripkean view of possible worlds. When he asks us to perform the thought-experiment of imagining our own zombie-twin, he is taking a stance on the issue of transworld identity. If I imagine my own zombie-twin, then I am imagining a possible world containing an individual which I need to be able to identify. This is the so-called problem of transworld identity and there are many positions people have taken in order to solve it. Dietrich and Gillies point out, however, that Chalmers believes the problem of transworld identity to be irrelevant to his enterprise. Thus, "we have that Chalmers is not a typical modal realist about possible worlds" (Dietrich & Gillies 2001, 368). If he were a typical modal realist, he would need to provide his position regarding transworld identity. But, they

argue, even if he's not a typical modal realist, Chalmers needs to be some sort of modal realist. Since Chalmers relies on transworld identity—he needs to be able to pick out his own zombie-twin—and does not provide his own solution to the problem of how to identify inhabitants of possible worlds, he must subscribe to a position very similar to Kripke's.

Kripke's position concerning possible worlds is that they are *stipulated*. Rather than there being concrete possible worlds containing possible individuals, possible worlds and their inhabitants are stipulated. Thus, there is no transworld identity problem at all, because when we stipulate that we are talking about Richard Nixon, we don't have to do any more than this. The stipulation is enough to establish that it is *Richard Nixon* whom we are discussing.

This is the kind of modest modal realism which Chalmers must hold, according to Dietrich and Gillies. As this is the case, then, "when we are imagining zombie twins, we are simply imagining counterfactual situations about ourselves in which we do not have qualitative experiences" (Dietrich & Gillies 2001, 370).

To proceed, we must consider Chalmers's notions of primary and secondary intensions. Both notions have to do with how the reference of a term is fixed. The notion of primary intension corresponds to the way a term's reference is fixed in the world where the term is being used. In Putnam's twin-earth experiment, the primary intension of "water" would be the liquid substance contained in lakes, rivers, streams on twin-earth. The substance would be XYZ. Secondary intension is what we have when the term's reference has been fixed already in the actual world, and the term is being used counterfactually. Thus, once the reference is fixed in the actual world for a certain term, such as "water," the secondary intension does not change. "Water" refers to H2O in all possible worlds.

According to Dietrich and Gillies, whenever we are dealing with primary intensions and possible world inhabitants, we have to flesh out what it is we are referring to in the possible world. “Water,” under its primary intension, on twin-earth would not be referring to H2O, but XYZ. But we would have to know how the reference is fixed on twin-earth, in order to know what it’s referring to. The way we would go about doing this, say the authors, would be to give something similar to a definite description.

Now, Chalmers also supposes a primary conceivability (to which we will return shortly) that corresponds to primary intension. Further, there is a primary possibility that corresponds to primary conceivability. Dietrich and Gillies contend that Chalmers must tie primary conceivability to primary possibility. But it is precisely this that he cannot do, given his distinction of primary and secondary intension and his commitment to a Kripkean position concerning transworld identity.

If it’s true that we are imagining actual-world individuals in the possible world, with the only difference being that the inhabitants of the possible world do not have phenomenal experience, then each of us may only imagine our own zombie-twin. Dietrich and Gillies note that Chalmers prizes the epistemic asymmetry of conscious phenomena, which implies that I cannot inspect some other person’s zombie-twin to see if it has consciousness. I can only inspect my own, since my zombie-twin is just me without consciousness.

The problem lies in how we flesh out the primary intension of the terms referring to individual zombie-twins. We may take Chalmers’s zombie-twin as an example. “From what we know about primary intensions it is unproblematic to say that the primary intension of ‘Chalmers’ is ‘chalmersy stuff’ … But what is chalmersy stuff, anyway? Saying ‘Chalmers is

chalmersy stuff' does not provide enough information about the primary intension of this concept to evaluate the possibility of the situations at hand" (Dietrich & Gillies 2001, 373). Chalmers would have to give something similar to a definite description, as we saw in the case of "water" above. What would serve as a definite description for "Chalmers?" "Now, it seems reasonable that whatever the description is that Chalmers uses here to flesh out the primary intension of 'Chalmers'—i.e., to give substance to 'chalmersy stuff'—it must have consciousness as a constituent member" (Dietrich and Gillies 2001, 373). Thus, it's not possible to flesh out the primary intension of "Chalmers's zombie-twin" without working in the notion of the zombie-twin's having consciousness.

Dietrich and Gillies conclude, then, that since "all of this was to evaluate the logical possibility of counterfactual situations about ourselves, it follows that our zombie twins are not [primarily possible]" (Dietrich and Gillies 2001, 374).

Someone might question their claim that consciousness is a constituent member of any description of Chalmers. They don't address this objection, though they do address the objection that a purely physical description would be adequate to flesh out the primary intension. Overall, their consideration of C-P, especially in light of Chalmers's own distinctions, seems to raise plausible reasons for rejecting the conditional statement.

Another attempt made against Chalmers's use of C-P is made by Joseph Levine (1998). Levine argues that modal intuitions are primarily "a sensitivity to the logical forms of the representations by which we conceive them" (Levine 1998, 450). Thus, he defines metaphysical possibility, in part, by stating the following, which he calls the Modal Bridge Principle (MBP):

(MBP) A situation *S* is metaphysically possible just in case it has no accurate representation that is logically inconsistent. (Levine 1998, 450)

So, the notion of possibility is extended only so far until we reach what we recognize as a logical inconsistency contained in an accurate representation. But having not arrived at a logical inconsistency, we cannot then assert that something or other is possible. "On this way of understanding the relation between possibility and conceivability, conceivability emerges as our only guide to possibility (for non-actual situations, that is), but not as a guarantee of possibility" (Levine 1998, 451). So, prima facie, C-P is not true. But it's not quite that simple, according to Levine, when it's applied to the mind-body problem.

The anti-materialist relies on what Levine calls the distinct property model (DPM). The DPM is a model which is an attempt to explain situations in which we know the meanings of two terms and yet must *empirically* discover that the two terms refer to the same object. Knowing the meanings of the terms alone is not sufficient for knowing that they refer to the same thing. "It is the distinctness of the properties through which we represent the objects/states in question, and the fact that these properties are only contingently related, that explains our inability to determine a priori that the objects/states are one and the same" (Levine 1998, 455).

This applies to the case of pain, according to the anti-materialist. Even if we grant the possibility of empirically discovering that pain just is C-fibers firing, we must also admit that the properties of pain and the properties of C-fibers firing will relate to one another only contingently. Thus, there is a possible world in which pain is not C-fibers firing, because the relevant contingent properties do not relate to one another in the possible world. Thus,

materialism is false.

The anti-materialist is committed to DPM, according to Levine. And if DPM is a good model, then the materialist is in trouble, because there would be some properties of pain which would be irreducibly mental. But there might be a way out for the materialist. Levine holds that the reference of a term is fixed by its "mode of presentation" and the contextual features of one's acquisition and use of the term. By "mode of presentation," he means "the means by which a representation connects to its referent" (Levine 1998, 457). He argues that we should take note of a particular distinction involved in the concept of mode of presentation—those modes he calls ascriptive and those he calls non-ascriptive. He cashes this distinction out thus: "An ascriptive mode is one that involves the ascription of properties to the referent, and it's (at least partly) by virtue of its instantiation of these properties that the object (or property) is the referent. A non-ascriptive mode is one that reaches its target, establishes a referential relation, by some other method" (Levine 1998, 457). Levine believes that, in order to avoid the analytic-synthetic distinction and meaning holism, it makes sense to take modes of presentation as non-ascriptive, in which case, the property that fixes the reference of a term need never be known to the user of the term at all. But if we accept that modes of presentation are non-ascriptive, rather than ascriptive, then, argues Levine, DPM should be revised. "We still appeal to distinct properties associated with the two modes of presentation, but we drop two crucial conditions: (1) they don't have to be contingently related, and that's because (2) they don't have to be explicitly represented within the mode" (Levine 1998, 458).

As this revised version now seems plausible, the anti-materialist argument may be seen to founder on its claim that conceiving of zombies implies their possibility. "Once we allow the

possibility of non-ascriptive modes of presentation for qualia, the inference from the conceptual possibility of zombies to their metaphysical possibility is undermined" (Levine 1998, 458). Since there is a logical inconsistency involved, it may be true that zombies are conceivable, but, according to Levine's characterization of metaphysical possibility, they may not be possible. The possibility of zombies is not entailed by the fact that one can conceive of them.

A slightly different approach to attacking C-P is taken by Peter Marton (1997). Marton charges not that C-P is false, but that it’s actually too strong for what the anti-materialist wants to accomplish. He begins by claiming that "zombists" are committed to an S5 semantic model. The S5 model is the one in which "relative possibility is taken to be an equivalence relation; that is, it is a relation that is reflexive, symmetrical, and transitive" (Loux 1979, 24). As this is the case, the anti-materialist conclusion which zombists push would lead to the following: If materialism is false, then it is necessarily false. Since the phenomenal does not supervene on the physical in the actual world and since the accessibility relations between the actual world and possible worlds are symmetric, the phenomenal will not supervene on the physical in any possible world. And this would lead to the statement that materialism is necessarily false.

But if materialism is necessarily false, then it is impossible. If materialism is impossible, then it is inconceivable, based on the contraposition of C-P itself. The zombist paints conceivability as cases in which the imagining of something does not contain explicit contradiction. Chalmers himself makes statements reflective of this characterization. Thus, the fact that materialism is inconceivable entails that there is an explicit contradiction involved in imagining its being true. The burden of proof, by the zombist's own admission, falls on one claiming that an imagined scenario contains an explicit contradiction. Therefore, the zombist

would have to produce proof of a contradiction's being involved in entertaining the likelihood of materialism's being true. Marton does not believe that this can adequately be done.

Thus, Marton does not argue that C-P is false or misapplied. It's actually too strong to accomplish the zombist's needs. The zombist does not want to put herself in the same predicament she puts the materialist in—having to produce contradictions in order to refute arguments. But she does just this, when she utilizes C-P. In the end, she is hanged by the same rope that she had cut for the materialist.

Now that we have seen that there are some rather plausible objections to C-P, I will provide two arguments of my own regarding whether conceivability is a good guide to possibility.

The Self-Undermining Feature of C-P

My first argument relies on the fact that, if C-P is true, then it is necessarily so. I will discuss why I think this is plausible shortly. For now, I will present my argument, assuming that C-P is necessarily true. If C-P is necessarily true, then its negation would be impossible. There would be no possible world in which C-P is false. A fortiori, the negation of C-P would be false in the actual world, as the actual world is just one of the possible worlds. But as we have just seen, there are several individuals (not including Marton) who believe strongly enough that C-P is false to argue for their belief. How could they *believe* C-P is false, unless they could conceive of its being false? There has to be some content to their belief, and the content would have to be fine-grained enough for them to be able to imagine its being false. They provide rich descriptions

of their beliefs in their respective arguments. But, if C-P is necessarily true, if true at all, and its negation is impossible, then what these people are conceiving would be impossible. If they are conceiving of C-P's being false, while the falsity of C-P would be impossible, then this would be a counterexample to C-P itself. Thus, far from being necessarily true, it would actually be false, at least in the actual world.

Now, I am not only relying on C-P's being necessarily true, if true at all; I am also assuming that at least some of the individuals who believe C-P is false are really conceiving of its being false. An opponent might object to this, however. If C-P is necessarily true, they might say, then how could anyone *possibly* conceive of its being false? They might argue that what I have demonstrated is not that conceiving of C-P's being false actually undermines C-P itself, but that putative conceivings of its being false are not conceivings at all. One person's modus ponens is another person's modus tollens. They might further argue that, before Kripke, some people thought they were conceiving of water's not being H2O. Now we know that, given its secondary intension, "water" necessarily refers to H2O. So people were not conceiving of water's not being H2O, but something that seemed to be water, but wasn't. The same thing might be happening here. The individuals whose arguments I outlined above may be conceiving of something they believe is C-P's being false. In fact, since C-P is necessarily true, they cannot be conceiving of this at all.

In order to respond to this objection, I will need a good definition of "conceivable." In the next section, I will consider possible definitions in the search for an adequate one. I will postpone my reply until then.

I still need to provide an argument for why I think Chalmers, and anyone using C-P,

would be committed to its being necessarily true. I won't be attempting an argument that establishes certainty on this matter. Rather, there will a lot of hand-waving and reasons why I believe my claim is plausible.

First, we should note that it is widely accepted by those who make conceivability arguments that if C-P is true, we should be able to know it a priori. If our knowing C-P depended on experience, then it might be the case that we just haven't observed its failure to hold. In other words, there could be cases in which conceivability does not imply possibility that we just never observe. If this were true, or even possible, however, then we could not move from epistemic possibility to metaphysical possibility. Our failure to know of the cases in which C-P fails would not ensure that there are none. Whether C-P is a priori or not I will not decide here. Suffice it to say that people who make arguments that utilize C-P generally accept that it is a priori. Chalmers definitely seems to take C-P as an a priori statement.

But does its being a priori imply that it is necessary?[6] We know, post-Kripke, that the concepts "a priori" and "necessary" do not completely overlap in the actual world. There are necessary a posteriori statements. There are a priori contingent statements. This doesn't call into question the fact that *most* a priori statements are necessarily true. What was debunked by Kripke was the claim that a priori statements just were necessarily true statements. And vice-versa. We don't have to deny that there is a close relation, however. It seems fairly uncontroversial that most a priori statements which we might make are necessarily true statements. And vice-versa. So, it is not *implausible* to suggest that if C-P is taken to be a priori, it should be taken as

[6] For a good discussion of this issue, see (Bonjour 1998).

necessarily true.[7]

Are there reasons for thinking that it is *plausible* that Chalmers is committed to C-P's being necessarily true? I think that, in relation to Chalmers's zombie-argument, there is reason to think that he depends on its being necessarily true. If C-P were a priori, but contingent, then it would not be the case that C-P is true in every possible world. In other words, there would be possible worlds in which C-P is false. Furthermore, there wouldn't be any conclusive reason for thinking that the actual world happens to be one of the worlds in which C-P is true. Chalmers cannot be satisfied with that. His strategy is to establish that zombies are possible in virtue of the fact that they are conceivable. If C-P were only contingently true, then it might be the case that the conditional is not true specifically in relation to the zombie-argument. We don't know whether zombies are possible without the aid of C-P. Thus, the consequent might be true or false. If C-P is only contingently true, then the fact that zombies are conceivable would not imply their being possible. At least, we wouldn't know this for a fact, unless we already knew whether zombies are possible. Each instance of the conceivability of an object leading to the possibility of the object would confirm our belief that C-P is true. But there could arise a counterexample at any moment. Thus, C-P in relation to the zombie-argument would just be undetermined. Chalmers might have the option of claiming that the conceivability of zombies makes their possibility highly likely. But he needs more than mere high likelihood.

So, it seems plausible, to me at least, that Chalmers is committed to the claim that C-P is necessarily true. If he does hold this, then he has to argue that anyone who would claim that C-P

[7] As I said above, I am not trying to decide whether C-P is a priori. I argue that Chalmers and those using C-P are committed to its being a priori. Taking that as established, I then proceed to argue that the fact that they are committed to its being a priori suggests that they would also be committed to its being necessarily true.

is not true would not be conceiving of C-P's being false. If they are conceiving of its being false, they are conceiving of an impossibility, which would actually refute C-P by itself. But it seems rather quick to make the flat-footed claim that anyone who opposes C-P is actually mistaking an inconceivable scenario for a conceivable one.

The Yablo-Chalmers Picture of Conceivability and Goldbach's Conjecture

My second argument against C-P will begin by our considering what Chalmers has in mind when he uses the word "conceivable." He relies much on (Yablo 1993), so I will sketch Yablo's discussion of the term and shift to Chalmers's discussion when the need arises.

Yablo begins by considering whether "conceivability" is equivalent to "believability." He notes that in daily conversation, we often conflate the two terms. In philosophical contexts, however, we hardly ever mean "believability" when we use the word "conceivability." For, he says, if conceivability entailed believability, "then whenever one was *certain* that something was not the case, one would be unable to conceive it even as a possibility!" (Yablo 1993, 11) And if believability entailed conceivability, then we would run into trouble whenever we came across cases in which we do not have enough evidence to conceive of a proposition being true, but we could perfectly well believe that the proposition is true. These would be cases in which the truth-value of a proposition is not yet known. Yablo uses as an example Solomon and a test that Sheba sets up for him. Before him are a real flower and a wax flower. He is then told to pick the real one, without smelling or touching it. Before Solomon enlists the aid of a garden bumblebee, he

does not know which flower is real. But he can perfectly well believe that the flower-like object before him, which, it turns out, is the artifact, could sprout flowers. Now, he does not have enough evidence to know which is which, so he is not conceiving of the artifact before him as sprouting flowers. He is undecided about which is the real flower. But this does not mean he cannot believe that it can sprout flowers. This is what Yablo calls an impure example, because the proposition stating that the object which turns out to be the artifact should sprout flowers is not positively inconceivable. Pure examples are much more difficult to produce.

Another argument that Yablo makes for this point is that, if we were to think of conceivability strictly as believability, and we tried to plug that conception directly into C-P, then we would actually get cases in which the less evidence we have for a proposition, the more likely it is that it's possible. This is not right, however. So conceivability should not be thought of merely as believability.[8]

Finally, Yablo arrives at what he believes is close to what we mean when we use the term "conceivable." He writes, "I find *p* conceivable if I can imagine, not a situation *in* which I truly believe that *p*, but one *of* which I truly believe that *p*" (Yablo 1993, 26). So, conceivability is to be understood as a certain kind of imagining. But what kind of imagining? He makes the distinction between two kinds of imagining: *propositional* and *objectual*. Propositional imagining is *imagining that*. Objectual imagining is *imagining perceptually*. Thus, I can imagine

[8] This might seem like a viable objection to my claim above that the fact that many people have objected to C-P indicates that they believe that C-P is false which, in turn, indicates that they can conceive of C-P's being false. Yablo's point, however, is not that believability *never* implies conceivability. He just wants to make it clear that the two concepts can be untied. This is not to say that there isn't a fairly constant co-instantiation of believability and conceivability. More to the point, however, is that *philosophers* would not haphazardly make objections which they believe to be good ones, without having a clear conception of what they are arguing. At least, ideally.

that today is Tuesday, and this would be propositional imagining. I can imagine a Tuesday spent in the Jardin du Luxembourg, replete with green grass, pigeons, sparrows, and lounging French people, and this would be objectual imagining. Oftentimes we do both. I can imagine, at the same time, that it's Tuesday and all of the rich details in the Jardin du Luxembourg. This would be a case in which I was imagining in both of the ways Yablo describes.

Taking these distinctions, Yablo then arrives at a more solid description of conceivability. "Conceiving that *p* is a way of imagining that *p*; it is imagining that *p* by imagining a world of which *p* is held to be a true description" (Yablo 1993, 29).

Now we must move from Yablo's discussion to Chalmers's, because Chalmers wants to make further distinctions concerning our use of the term "conceivable". The purpose of making these further distinctions is clear: "By making these distinctions, I think at least one plausible and defensible conceivability-possibility thesis can be formulated, free of any clear counter-examples," (Chalmers 2002, 146).[9]

The first distinction is between prima facie and ideal conceivability. Something is prima facie conceivable when it is conceivable on first appearances. "That is, after some consideration, the subject finds that *S* passes the tests that are criterial for conceivability" (Chalmers 2002, 147). Something is ideally conceivable when it's still conceivable, even after ideal rational reflection. Chalmers cashes out ideal rational reflection as having "justification that cannot be rationally defeated" (Chalmers 2002, 148). Often, things will be merely prima facie conceivable without

[9] This clearly indicates that Chalmers believes that for his conceivability argument to succeed, C-P must not fall prey to counterexample. His providing these new distinctions is an attempt to guard against counterexample. Thus, if a counterexample is offered up and C-P, in whatever form Chalmers gives it, cannot stand against it, C-P fails.

being ideally conceivable. Thus, a proposition could be prima facie conceivable, but not ideally conceivable.

The second distinction is between positive and negative conceivability. Negative conceivability occurs when "*S* is not ruled out a priori, or when there is no (apparent) contradiction in *S*" (Chalmers 2002, 149). Positive conceivability occurs when one can form a positive conception of the situation in which *S* is the case. Now, according to Chalmers, positive conceivability is similar to what Yablo describes when he talks about propositional and objectual imagining. Chalmers wants to add one more kind of imagining to the positive sense of conceivability—*modal imagining*. Modal imagining is, in a sense, a combination of objectual and propositional imagining. He writes, "[I]n this case, we have an intuition of (or as of) a *world* in which *S*, or at least of (or as of) a situation in which *S*, where a situation is (roughly) a configuration of objects and properties within a world. We might say that in these cases, one can *modally imagine* that *P*. One modally imagines that *P* if one modally imagines a world that verifies *P*, or a situation that verifies *P*" (Chalmers 2002, 151). It should be clear that this is, in a nutshell, Yablo's definition of "conceivable." Thus, Yablo's conception is only one particular kind of conceivability, according to Chalmers.

The third distinction is between primary and secondary conceivability. Primary conceivability occurs "when it is conceivable that S is *actually* the case." Secondary conceivability occurs "when S conceivably *might have been* the case" (Chalmers 2002, 157). These distinctions correlate roughly to Chalmers's notions of primary and secondary intension, outlined above.

Now we are in a position to answer the objection proposed earlier to my claim that I am actually conceiving of C-P being false, when it would be necessarily true, if true at all. Chalmers might argue that I am engaged in prima facie positive primary conceiving when I claim that I can conceive of C-P's being false. Consequently, my conceiving of C-P's being false would not imply that it's possible that C-P is false. But I cannot find a contradiction, even upon rational reflection. I can imagine necessary statements being false. After all, I can imagine impossible states of affairs. I think I can imagine God creating a stone too heavy for himself to lift. He waves his hands, a stone appears; he then stoops to pick it up, groans, grabs at his back, and shakes his head. Why then could I not imagine C-P's being false? I can imagine a situation in which someone conceives of zombies, and then finds out later that there is a logical contradiction involved in hypothesizing zombies. I simply cannot find any contradiction in imagining C-P's being false. Thus, it would seem that my conceiving would actually be ideal. If so, then the objection made earlier really doesn't weaken my argument.

So how do these distinctions affect our understanding of the zombie-argument? According to Chalmers, the kind of conceivability we are talking about in relation to zombies is ideal, primary, positive conceivability. And this kind of conceivability, he contends, does imply possibility.[10] Prima facie conceivability, he admits, is not a good guide to possibility. Thus, the strategy one might adopt in order to attack the zombie-argument would be to try to show that

[10] It might be argued that this claim cannot be defeated, because ideal rational justification can simply be extended out as far as one wishes. Since the justification is deemed *ideal*, then, if zombies are possible, *ideal* reflection probably would bear out the possibility. To claim that zombies are ideally conceivable would necessarily imply that zombies are possible. Then the question arises as to whether zombies are ideally conceivable. (We will answer this question shortly.) If this is how Chalmers wants the story to go, however, then his statement that there is one kind of conceivability-possibility thesis that is not open to counter-example is seen to be empty. Of course, if we define conceivability in this way, possibility will necessarily follow! We might object then that defining conceivability like this doesn't answer the original question, which had to do with conceivability in the ordinary, unrefined sense.

zombies are not ideally conceivable. This would present a problem, however. He writes, "[D]efeating ideal conceivability will require an a priori entailment from physical to phenomenal, which requires that phenomenal concepts can support that entailment" (Chalmers 2002, 197). He doesn't believe that this can be done.

Let's return, for a moment, to a case which Chalmers believes is only prima facie positively conceivable—the case involving Goldbach's Conjecture (GC). If GC is true, then it is necessarily true. If it is false, then it is necessarily false. Thus, if I can conceive of its being true, when it is actually false, or conceive of its being false, when it is actually true, then I can conceive of its being true and its being false, even when it is only one or the other. Even if I don't know whether GC is true or false, the fact that I can conceive of its being true and I can conceive of its being false implies that I am conceiving of something that is impossible.[11] If this is the case, then we have a counterexample to C-P.

Now, according to Chalmers, what I am doing when I think I am conceiving of GC is imagining there being a lot of evidence for GC's being true (or false). He writes, "The case where one conceives of mathematicians announcing a proof of Goldbach's conjecture (or its negation) is best seen as a case where a superficial prima facie positive conceivability judgment is undermined by a moment's reflection" (Chalmers 2002, 160). When we reflect on what we are imagining, we find that we are not imagining GC's being true (or false), but only a situation in which there is justification for accepting either.

[11] Of course, one would have to accept the principle of bivalence, too, for if GC were supposed to be neither true nor false until we discover its proof or refutation, then I would be merely imagining a proposition as being true when in fact it is neither true nor false. But if GC is neither true not false, then it certainly could not be necessarily true or necessarily false, at the moment when I imagine its being true or false. I take it, however, that the principle of bivalence is still widely accepted.

According to Yablo, the problem is in the detail which we are imagining. Let's say that when we conceive of GC's being true, we are imagining a computer spitting out a proof. And when we conceive of GC's being false, we are imagining a computer spitting out a counterexample. Yablo then asks, if we don't know whether GC is true or false, how can we know that the proof (or counterexample) is correct? "Am I to *imagine* it set out in convincing detail? But if the detail is only *imagined* to be convincing, it does nothing to increase my *actual* confidence in the proof's correctness?" (Yablo 1993, 32) But, if the detail that we need is actually imagined, then we have simultaneously imagined the actual proof (or counterexample). If that were the case, then the proof (or counterexample) is there to be produced and the person imagining GC's being true (or false) has actually proved GC true (or false). What this shows, however, according to Yablo, is that GC is not inconceivable or conceivable, but undecidable.

Of course, Chalmers wishes to say that it's merely prima facie conceivable, not undecidable. We should note here how the evaluation of GC relies on the grain of detail which is required. When we imagine a bunch of mathematicians announcing that GC has been proven, this is too coarse-grained. Even when we imagine seeing a proof of GC, say from a computer output, this is still not fine-grained enough. In fact, it might be argued, we will never be able to conceive of GC's being true (or false) in the way that Yablo and Chalmers would accept, until GC is proved true (or false). Presumably, *someone* would have conceived of it before that—the person or group of persons who discover the proof. But, according to Yablo's and Chalmers's description, it would have to be a simultaneous conceivability and discovery.

So, let's accept that the case of GC is prima facie conceivable, primarily because the grain-level of imagining which would be required in order for us to conceive of its being true (or false) is simply not possible, without knowing beforehand whether GC is true or false.

What about the grain-level required in order to imagine zombies? Is the situation any different between the zombie-argument and the GC-case? According to Chalmers, the only way one could demonstrate that zombies are not ideally conceivable is to show an a priori entailment from the physical to the phenomenal. Chalmers contends that demonstrating this entailment from the physical to the phenomenal cannot be accomplished, while it *has been shown,* via the zombie-argument, that the a priori entailment does not hold. The zombie-argument was intended to show that the a priori entailment does not hold; turning to this fact, however, in order to buttress the zombie-argument is circular. The best that can be said here would be that the a priori entailment has not been established. But it hasn't been refuted, either.

Chalmers puts the cart before the horse. His zombie-argument is intended to show that since zombies are conceivable, they are possible. That zombies are ideally conceivable, however has not been established. The zombie-argument is supposed to show that science is wasting its time as it operates on the assumption that consciousness may be explained in physical terms, but it seems more likely that science will have to determine whether consciousness may be explained in physical terms before we can determine whether zombies are ideally conceivable.

Chalmers seems to rely on the mysteriousness of consciousness I mentioned at the beginning, in order to bolster his claim that ideal reflection does not produce a contradiction. The fact is—we don't know whether the physical entails the phenomenal. But science appears to be prepared to eventually give us this answer.

Similarly, GC has not been proved or disproved, and we don't know which proposition—that GC is true or that GC is false—is necessarily true and which is necessarily false. The truth-value of GC is undetermined. We don't reason from our being able to conceive of GC's being true to the assertion that it's possible for GC to be true, for the simple reason that, if GC is false, then it's necessarily false. We wait for the mathematical world to tell us whether GC is true or false. And if they never tell us, we take that fact as evidence that we will never know which of the two propositions is ideally conceivable.

I contend that the zombie-argument is relevantly similar to the GC-case. In the GC case, one of the scenarios is coherent, though we don't know which one. In the zombie-case, we don't know whether the physical entails the phenomenal. If it doesn't, then zombies are possible. If it does, then they aren't. But, as this is the case, how can we call zombies positively ideally conceivable and not call the scenarios of GC ideally conceivable? Zombie-arguments are in the same boat as GC scenarios. We don't want to say that both GC scenarios are ideally conceivable. Consequently, zombies are prima facie conceivable. As Chalmers admits that prima facie conceivability is not a good guide to possibility, his argument does not succeed.

I hope that it is clear that I am not really calling into question whether zombies are conceivable, at least, in the context of Chalmers's distinctions. I think that we can conceive of zombies at a coarse-grained level. But, by Chalmers's own admission, this kind of conceivability does not imply possibility. So, I don't think that the first premise, "If zombies are conceivable, then zombies are possible," is true.

To recap the section: there are several arguments against C-P. I present two arguments of my own. The first points out that, if C-P is necessarily true, if true at all, then it actually

undermines itself, for it is possible to doubt whether C-P is true. And to doubt whether C-P is true is to conceive of its being false. But if it is necessarily true, if true at all, then one's conceiving of its being false serves as a counterexample. Thus, C-P is either false or not necessarily true. I argue that, in order for C-P to do the metaphysical work Chalmers intends it to do, he is committed to C-P's being necessarily true. Thus, C-P is false.

My second argument relies on Chalmers's distinction between prima facie and ideal conceivability. He believes that prima facie conceivability is not a good guide to possibility. Further, he argues that the zombie-argument is ideally conceivable. I argue that the zombie-argument is relevantly similar to cases where we can conceive of Goldbach's conjecture being true, when it's actually false, or where we can conceive of its being false, when it's actually true. Chalmers contends that Goldbach's conjecture cases are only prima facie conceivable. Thus, I conclude that the zombie-argument is only prima facie conceivable. Consequently, the conceivability of zombies does not lead to their being possible.

There are some compelling arguments which attempt to refute Chalmers's claim that zombies are conceivable. We turn now to those.

Conceivability of Zombies

Even if it's true that conceivability implies possibility, Chalmers's argument is in trouble if zombies are, in fact, not conceivable. We will look at just a few attempts which have been made to show that zombies are inconceivable.

Daniel Dennett writes, "[W]hen philosophers claim that zombies are conceivable, they invariably underestimate the task of conception (or imagination), and end up imagining something that violates their own definition" (Dennett 1995, 322). Dennett goes on to consider what he calls a "zimbo", which is defined as "a zombie equipped for higher-order reflective informational states." In other words, the zimbo has a particular faculty that acts functionally as consciousness is supposed to. Then, he asks the believer in zombies whether these zimboes are conscious, after all, or whether the thing they hypothesize is missing in zombies is even relevant to the mind-body problem at all. His objection to zombies is along functionalist lines. Can we fully imagine zombies as behaviorally, functionally, and psychologically identical to people of the actual world? Or are we imagining zimboes, which do possess something which functions like consciousness?

A more specific objection to the conceivability-claim can be found in a paper by Allin Cottrell. He asks *what is it* that we are supposed to be imagining when we imagine a zombie? The instructions Chalmers gives us—imagine something physically identical to an actual world human, but dark inside—is not clear. The best conception we might produce, according to Cottrell, is the following: "A zombie, then, is like a person on the outside, but just like a rock on the 'inside'" (Cottrell 1999, 7). But the problem occurs when we try to tie these two facts together in a coherent whole. "The task is somehow to hold in our heads simultaneously the notion of normal human behavior, on the one hand, and internal nullity, on the other" (Cottrell 1999, 8).

Normally, we are able to tie concepts together. Someone might ask us to imagine unicorns. We would do so by imagining a horse and imagining a horn and combining the horse

and horn into a coherent imagined creature. We could then "see" the creature we have imagined and verify that both the concept of "horseness" and "hornedness" are evident in the picture. But can we do the same thing in the case of zombies? Cottrell says no.

> In the first two cases considered above, we concluded that the burden of proof would rest with anyone claiming that the putative conception (unicorn, sawdust man) was not properly conceivable. One reason for this conclusion was that in those cases it was possible to construct an external 'verification' of the consistency of the conception. … But note that no such verification is possible, even in principle, in the case of the zombie. (Cottrell 1999, 8)

Now, it might be objected to Cottrell's description above that he is attending to the objectual imagination only. If he were to utilize the notion of propositional imagining, then he would see that imagining normal human behavior could be done objectually, while imagining internal darkness would be propositional, i.e., imagining that there is nothing going on phenomenally inside. There would seemingly be no problem, until one were to ask what is meant by the proposition, "There is nothing going on phenomenally inside." In other words, part of conceiving of the zombie would be *understanding* the proposition which would accompany the imagining. But, in order to understand the proposition in the present context, one would have to know *how* the normal human behavior could be instantiated while there is nothing phenomenal going on inside. Thus, the problem Cottrell points out would still exist.

An even more persuasive argument against the conceivability of zombies can be found in the paper "Why Zombies are Inconceivable" by Eric Marcus. Marcus attacks Chalmers on his own turf, as it were, showing how zombies are not primarily positively ideally conceivable. Marcus begins his attack by distinguishing between *third-person* differences and *first-person* differences. He explains what he means by these terms in the following: "To imagine Abe Lincoln third-personally is to imagine, for example, the way his parts are laid out in space. To imagine him first-personally is to consider how the world appeared to him, to imagine what he felt, experiencing his moods, and so forth" (Marcus 2004, 56). According to Marcus, the difference we are asked to imagine in conceiving of zombies is first-personal.

In the zombie thought-experiment, the difference is between the presence of consciousness in actual world beings and the absence of consciousness in the zombies. Marcus asks, however, "But what is it to imagine first-person absence?" Third-person absence is possible. We have a background of presence to measure it against. But first-person absence is not analogous. "There is no inner border or background of inner space against which it is possible to conceive subjective absence. Imagining subjective absence presents an insurmountable obstacle. On the one hand, it is something that we are to imagine from the first-person point of view … On the other hand, there is nothing that it's like to be subjectively absent" (Marcus 2004, 57).

So, if we are to do as we are asked, we would have to imagine ourselves experiencing *nothing*. But this seems self-contradictory. The problem, according to Marcus, is that imagining zombies requires imagining nothing. So, "to 'imagine' a zombie is not really to imagine at all" (Marcus 2004, 57). Thus, upon ideal reflection, we discover that zombies really aren't

conceivable. They were, however, apparently conceivable. So, they were prima facie conceivable. But what deceives so many into thinking that they are conceivable?

Marcus argues that "those who take themselves to imagine zombies are mistaking not imagining something for imagining nothing" (Marcus 2004, 58). They think they are imagining a lack of consciousness when actually they just aren't imagining consciousness. The former cuts some metaphysical ice, as Marcus puts it. The latter does not.

Thus, Marcus concludes the following: "Although there is a sense in which we can imagine a physical and functional duplicate of our world without consciousness, 'without consciousness' modifies the *act,* rather than the *object* of imagining" (Marcus 2004, 48). We could imagine a molecule-for-molecule duplicate of the actual world without imagining mountains, planets, plants, etc. But we cannot imagine a molecule-for-molecule duplicate of the actual world which does not contain mountains, planets, plants, etc. The latter is simply not possible. Thus, if we imagine a zombie-world, we are imagining a world physically identical to our own without imagining consciousness being a part of it. This does not imply that consciousness is non-physical.

The above arguments—especially Marcus's—seem to me compelling. While I myself do not argue that zombies are not conceivable, I do have the same intuitions as those above, evident in my claim that zombies are only prima facie conceivable. Thus, I find Chalmers's positive argument for zombies lacking, and weakened by the above arguments. So, if there is any hope for the zombie-argument, it will have to reside in the version that I am calling the burden-of-proof argument. It is to this that we now turn.

The Burden-of-Proof Argument

We began our discussion of the zombie-argument by noting that Chalmers really makes two arguments for zombies: the positive argument and the burden-of-proof argument. These two arguments are significant for many reasons, but primarily for their conclusions—the statement that zombies are logically possible. For reference, I will include Chalmers's argument against materialism again, so that we can see that the conclusion regarding zombies is present as premise 2:

1. In our world, there are conscious experiences.
2. There is a logically possible world physically identical to ours, in which the positive facts about consciousness in our world do not hold.
3. Therefore, facts about consciousness are further facts about our world, over and above the physical facts.
4. So, materialism is false. (Chalmers 1996, 123)

At this point, we should take note of an important difference between the positive and burden-of-proof arguments. We considered several objections to the positive argument. But they attempted to show not that the conclusion is false, but that it's not established by the premises. The claim that zombies are possible is consistent with the falsity of both C-P and the conceivability-claim. The positive argument can fail, while the argument against materialism

succeeds. How so? Premise 2 can be true even if the positive argument does not succeed. So a refutation simpliciter of premise 2 will not be achieved just by refuting the positive argument. But if we were to refute the burden-of-proof argument, premise 2 would be irrefutably false. Of course, there are at least two ways one can refute a burden-of-proof argument: show how the burden of proof is misplaced or show how the argument itself just does not work. I will be attempting the second way.

Recall that Chalmers has set up the argument against materialism in the following way: None of the premises alone establishes the anti-materialist conclusion. They work in concert, so that, if they are all true, the conclusion is supposed to logically follow. Premise 1 is not controversial. Premise 2 supports premise 3, so if we want to know whether premise 3 is true, we will have to determine whether premise 2 is true. Thus, we must determine whether premise 2 is true or not, by Chalmers's own design. He obviously counts on the positive argument for zombies to demonstrate that premise 2 is true. We have seen that the soundness of the positive argument is doubtful. So, as a final effort, Chalmers would have to rely on the burden-of-proof argument, which states that zombies are logically possible until proven otherwise. The burden of proof is on the person who claims that zombies are not logically possible.

So we have to keep in mind two things here: (1) Chalmers places the burden of proof on an opponent, so that zombies are possible unless proven impossible. (2) Chalmers relies on his audience assenting to each and every premise in order to establish the anti-materialist conclusion. The intended audience—primarily consisting of agnostics and open-minded materialists—is not going to merely assent. Most members of the intended audience willing to entertain Chalmers's argument against materialism are going to attempt to do just what Chalmers indicates must be

done: they are going to try and prove that zombies are not logically possible. Some, but not all, will agree, having not proven zombies impossible, that zombies may be assumed possible. But very few are going to accept that zombies are possible, without conducting some research of their own.

What sort of research will be needed? Presumably one would want to investigate whether there is a contradiction involved in hypothesizing the existence of zombies. What sort of contradiction would one seek? I contend that, despite what Chalmers says about *conceptual* contradictions being required, *logical* contradictions are what an opponent would have to produce. If someone were to provide a logical contradiction associated with the postulation of zombies, then zombies would be incontrovertibly impossible, and the burden-of-proof argument would fail. It is not so clear that the burden-of-proof argument would fail as a result of a demonstration of a conceptual contradiction. After all, there are countless propositions and states of affairs that are logically possible, but not conceivable.

How is the search for the contradiction to be conducted? Presumably, one would perform a *reductio ad absurdum*—assume that zombies are possible and then show how this assumption leads to a contradiction. The next step would be to figure out what the contradiction would look like. It can't be just any contradiction that we seek. While it may be the case that our discovering one logical contradiction entails all contradictions, the initial contradiction cannot be placed in the context of the physical world, which serves as background in possible world *Z*. In order to entertain Chalmers's thought-experiment, we have to stipulate that the physical facts in world *Z* are identical to the physical facts of the actual world. If we happen to produce a logical contradiction, then it does not matter whether the stipulation is then violated. But we cannot

violate the stipulation upfront, by allowing in contradictions that are unrelated to the postulated zombie-twin. Thus, the one contradiction we *should seek* cannot be one that is unrelated to the zombie-twin, even if, after having produced the contradiction, all contradictions follow. This is all in line with keeping to the terms that we accept when we agree to perform Chalmers's experiment.

So, let's see how it goes then, trying to prove zombies logically impossible, in order to decide whether we should assent to premise 2 above. Let's assume that zombies are possible. This allows us to posit the existence of the zombie-twin of an individual of the actual world. Chalmers seems to be quite attached to and protective of his own zombie-twin, so let's posit the existence of another. Take the zombie-twin of my cat, Louie.[12] In possible world *Z* it is true that *p*, where *p* is the statement, "There is a creature physically identical to Louie that lacks consciousness." The statement *p* would have the form

(1) $\exists x(Ix \,\&\, Lx)$; (There exists an *x* such that *x* is physically identical to Louie and lacks consciousness.)

Statement (1) seems to be the most obvious and appropriate statement that is true in *Z* and not true in the actual world. Having assumed that zombies are possible, and having made the claim that there exists a zombie in *Z*, we would then need to try to show that in world *Z* there is a

[12] We may assume for the sake of argument, contra Descartes, that animals are conscious beings, and, also, that the kind of consciousness we are discussing is relevantly similar to Louie's.

logical contradiction. Since (1) is the statement we have determined to be the most relevant to *Z*, we would then try to prove the negation of (1). This would be the statement

(2) $\sim\exists x\ (Ix\ \&\ Lx)$; (There does not exist an *x* such that *x* is physically identical to Louie and lacks consciousness.)

The most probable way one would arrive at (2) would be by showing that zombies *cannot exist* in *Z*. Whether this would be best accomplished by relying on actual world science or some other method we won't try to decide here.

Statement (2) is logically equivalent to

(3) $\forall x(Ix \supset \sim Lx)$; (For all *x*, if *x* is physically identical to Louie, then it is not the case that *x* lacks consciousness.)

Thus, we would need to prove this statement, or a logically equivalent version of it, in order to show that our original assumption leads to a contradiction.

What exactly does proving (3) involve? I.e., what would someone wishing to prove statement (3) have to demonstrate? Basically, (3) just says that being physically identical to

Louie is sufficient to produce consciousness. Thus, to prove (3), one would have to prove that nothing that has the physical make-up identical to Louie's could fail to possess consciousness. One main thesis of materialism states that the physical structure of the brain is sufficient to produce consciousness. In fact, the zombie-argument is an attempt to refute this thesis. But (3) seems to be a restatement of this thesis. The upshot, then, is that one who wishes to prove (3), thereby providing a negation of (1) and a logical contradiction, would have to prove a central claim of materialism. Since Chalmers's argument against materialism relies on its premises being true and forcing the conclusion, the only way to halt the argument's progress toward its conclusion would be to prove this central claim of materialism true. In other words, the only way to prove Premise 2 false is to prove the conclusion false. But, Chalmers's intention was to get us to assent to Premise 2, giving us no choice but to accept the conclusion. Therefore, since we are supposed to assent to the second premise before assenting to the conclusion, and furthermore, since we have to assent to the conclusion before we assent to the second premise (because of the above demonstration), we never arrive at the conclusion ... or we arrive at it before we begin. In other words, Premise 2 does not support the conclusion at all, or it begs the question.

Chalmers is faced with a dilemma. Either premise 2 does not support the conclusion at all, because the truth-value of the conclusion would have to be determined before the truth-value of premise 2 could be determined; or premise 2 begs the question against materialism, because in order to ascribe a truth-value to premise 2, we would have to assume that a central thesis of materialism is false. Let's assume that premise 2 does support the conclusion, in order to see how the argument against materialism would be circular. If premise 2 supports the conclusion, then we must be able to determine that it is true. But as we have seen, in order to determine that

premise 2 is true, we would have to ensure that premise 2 cannot be falsified. To falsify premise 2, we would have to establish that materialism is true. So the only way for us to establish premise 2 as true, in light of the positive argument's not working, would be to demonstrate that we are unable to prove false the anti-materialist conclusion. Thus, a central thesis of materialism has to be assumed false in order for premise 2 to be judged true. But the conclusion is supposed to follow from the premises, not the other way round. Therefore, if we determine that premise 2 is true, as we would have to do if premise 2 supports the conclusion, then we would have to first establish that the anti-materialist conclusion is true. Thus, if premise 2 supports the conclusion, then the entire argument is circular.

If Chalmers wants his argument to be non-circular, then he will have to admit that premise 2 does not support the conclusion. But premise 2 is crucial to the establishing of the conclusion. Thus, his argument against materialism fails, our having shown that the positive argument does not succeed and the burden-of-proof argument only leads to the dilemma we have presented above.

My point in this section has been an epistemological one: how can we know that premise 2 is true? But Chalmers has framed his argument such that we would have to know whether premise 2 is true in order to see that the anti-materialist conclusion has been established. Since the positive argument for premise 2 does not succeed, Chalmers has only the burden-of-proof argument to appeal to. This argument does not work, either, because when we begin to investigate whether premise 2 is true, we see that there is no way to determine that zombies are logically possible without first conceding that materialism is false.

So, I hope it is clear that, in spite of the fact that I am making an epistemological, rather

than metaphysical, point regarding the zombie-argument and its place in the argument against materialism, the point directly impacts Chalmers's argument. He sets up the burden-of-proof argument in such a way that it depends on his audience believing that each premise is true. Once the investigation begins, however, it becomes clear that the premise does not support the conclusion. In fact, it's the other way round.

There are problems with Chalmers's positive argument for zombies. There are problems with Chalmers's argument for zombies. At least, so I have argued. One might think that I would be content in leaving the situation as it is, concluding that Chalmers's zombie-argument, and his argument against materialism, which depends on the zombie-argument, are hopelessly flawed. I intend to raise one more objection to Chalmers's zombie-argument. I will argue that the zombie thought-experiment is misguided from the very beginning. Work done on consciousness will not be furthered by zombie-arguments or any similar thought-experiments. Indeed, they only serve to consume time which would be better spent attending to more relevant things.

The Inherent Uselessness of the Zombie-Argument

I will begin this section by providing an analogy to Chalmers's zombie-argument/argument against materialism. It will be relevantly similar to Chalmers's arguments, but relevantly different, as well.

I call the following the "dullard-argument". It goes like this:

(1) Imagine a world physically identical to the actual world whose inhabitants are physically, behaviorally, and functionally identical to the inhabitants of the actual world. They work, they play, they sleep, they have families, etc. Anyone observing these creatures going about their daily lives would be unable to detect any difference between them and their actual world counterparts.

(2) Further imagine that, in spite of the fact that these possible-world inhabitants are behavioral isomorphs of the inhabitants of the actual world, they never experience any enthusiasm. They are dullards. Possible-world Dad takes possible-world son to the Red Sox game—son talks about getting excited, fidgets in the car seat on the way to the game, smiles continuously. If you ask him whether he is enthusiastic about going to the game, he will reply, "Sure am." But, in fact, he is not experiencing enthusiasm, because enthusiasm does not exist in this world.

(3) It seems that I can imagine this, without any apparent contradiction.

(4) If I can imagine this, then it is possible.

(5) Therefore, there is a possible world in which everything is physically the same as the actual world, but where something is missing in the possible world.

(6) Therefore, materialism is false.

This is a compressed analogue of Chalmers's zombie-argument/argument against materialism.

Now, the dullard-argument seems to me to be analogous in certain, significant ways to the zombie-argument. We are asked to postulate a world physically identical to the actual world. Then, we are to stipulate that something that exists in the actual world does not exist in the

possible world. Further, we are then compelled to admit that the actual world contains something which the possible world does not. Finally, we are to conclude that materialism is false.

It might be argued that the dullard-argument works so closely to the way the zombie-argument works that it also is an effective way to refute materialism. But there are some relevant differences between the dullard-argument and the zombie-argument, as well.

First, enthusiasm is not a state of consciousness. How can this be proven? Let's assume for the moment that zombies are conceivable. We can imagine a world in which every human-like creature is physically identical to the humans of the actual world and none of the human-like creatures possess consciousness. Moreover, let's imagine a particular zombie—my zombie-twin—sitting on his couch one afternoon, watching TV, and eating potato chips. Suddenly, the zombie leaps up, brushes crumbs off his chest, and announces loudly to anyone who is around, "Today, I resolve to get back into shape!" He throws away all his junk food, enrolls in a health club, begins jogging regularly, completely changes his diet. Whenever he has any free time, he engages in some aerobic activity. What's worse, he begins to beleaguer his friends and family with admonitions regarding their own lifestyles. One day, they cannot take it anymore, and they confront him, telling him that if he doesn't moderate his behavior resulting from his newly adopted way of life, he will soon be healthy, but all alone.

Would we not describe the zombie's behavior as enthusiastic? I think we would ascribe enthusiasm to the zombie. But if the zombie exemplifies enthusiasm, and, by definition, does not possess consciousness, then enthusiasm is not a phenomenal state. Therefore, even if the dullard-argument were to be accepted as sound, it would not establish that consciousness is non-physical. At best, it would establish that there is something that is non-physical.

This brings us to the second difference between the dullard-argument and the zombie-argument. If we can describe even a putative zombie as a creature displaying enthusiasm, then it's fairly clear that we typically mark enthusiasm by behavior and not by some underlying mental state. If this is the case, then it's also the case that whenever behavior that typically marks enthusiasm is present, we have no reason to believe that enthusiasm is not present. Even if we observed someone exhibiting typically enthusiastic behavior and we received from the person sincere denials of an underlying enthusiasm, I think we would still describe the behavior as enthusiastic and might even question whether the person truly understands the word "enthusiasm."

So, to return to the dullard-argument, we recall that behavior that seems enthusiastic is present in the possible world. If what I have said above regarding enthusiasm is correct, then the possible world does contain enthusiasm, in spite of the fact that I mistakenly thought that I was able to conceive of a world identical to our own in which enthusiasm is missing. Thus, the conclusion is not established.

We know that it makes no sense to imagine enthusiastic behavior without imagining enthusiasm. We know beforehand that there is a conceptual contradiction involved. This is the crucial difference between the dullard-argument and the zombie-argument. As I noted before, we don't know whether there is a conceptual contradiction involved in hypothesizing zombies. This is why it seems immediately absurd when we confront the dullard-argument and does not seem so absurd when we confront the zombie-argument.

Now, Chalmers might respond by pointing out the fact that I have just shown that my thought-experiment involves prima facie conceivability, not ideal conceivability. After some

reflection, I am able to see that I cannot, in fact, coherently imagine a world physically the same as the actual world, but containing no enthusiasm. Thus, this is no analogy to the zombie-argument. We should recall the point made earlier, however, that Chalmers depends on the mysteriousness of consciousness to strengthen his claim that zombies are ideally conceivable. In the dullard-argument, it seems pretty obvious that imagining the possible world is not coherent. We don't know enough about the nature of conscious experience to judge whether the zombie-argument is ideally conceivable or not. This doesn't weaken my claim, however, that the dullard-argument is a viable analogue of the zombie-argument.

Let's look at a similar attempt to show how the zombie-argument is flawed. Dennett (1995) sets up a comparison with the concept "health." We can perform a thought-experiment similar to the zombie-argument and the dullard-argument. First, we are asked to imagine a world a portion of whose inhabitants engage in activities which would require robust health in the actual world. Then, we are to make the stipulation that these people are not healthy, according to the thought-experiment. Thus, there seems to be a mystery. They can do things like climb Mt. Everest, swim the English Channel, and yet not be healthy. Of course, Dennett does not think that there is a mystery at all—here or in the zombie-argument. He writes, "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" (Dennett 1995, 325). But, Dennett claims, health just is the proper working of all the parts of the body. Consciousness works the same way, he argues. "Supposing that by an act of stipulative imagination you can remove consciousness while leaving all cognitive systems intact … is like supposing that by an act of stipulative imagination, you can remove health while leaving all

bodily functions and powers intact. … Health isn't that sort of thing, and neither is consciousness" (Dennett 1995, 325).

Dennett makes the good point that consciousness is the wrong kind of thing to stipulate out of existence. His answer, however, is to define it functionally, as we do with health. I will not be adopting a functionalist position, but rather a deflationist position. I put forth a deflationist position, not because I don't find any mystery associated with consciousness, but for quite the opposite reason. I believe that the mystery of consciousness paralyzes us in some ways, so that we are afraid to proceed with the relevant research. And this paralysis manifests itself in empty activities, like thought-experiments that serve no purpose other than to keep the field preoccupied with questions of ontology. I recognize that there is still an explanatory gap between experience itself and the underlying mental processes of the brain. I further recognize that conceptual work is undertaken, in part, in order to close the gap. But if we have to greatly amend our conception of consciousness—as we would have to do if we are to consider consciousness as non-physical—as a result of these thought-experiments, then, I would argue, the gap doesn't get smaller, but larger, and we waste time in making the problem more intractable by bad research.

A Deflationist Description of "Consciousness" and its Relation to the Zombie-Argument

In addition to the problem I have just discussed regarding the analogy of enthusiasm, I now will argue, first, that current philosophy of consciousness needs a cohesive, agreed-upon description of the concept "consciousness;" and second, that the kind of description which is called for will not be such as to allow the zombie-argument to proceed as Chalmers would like it

to.

There are several levels of disagreement in contemporary philosophy of consciousness. First, there is disagreement regarding which experiences we should refer to as conscious. When one uses the phrase "conscious experience," it's not always clear what one means. Some philosophers conflate the terms "conscious" and "aware," so that one's experience of a shirt collar's rubbing the back of the neck would not count as conscious, if one were not aware of it at anytime. Also, when one is driving late at night and suddenly realizes that she has driven several miles without being aware of it is a case in which the driver was not conscious of driving. David Armstrong calls into question whether we really want to ascribe to drivers the ability to drive without being conscious of driving, at least at some level. Other philosophers make the distinction between being conscious and being aware. Ned Block distinguishes between phenomenal consciousness (or P-consciousness) and access-consciousness (or A-consciousness). P-consciousness is the qualitative experience of, say, seeing a sunset. A state is A-conscious if it is open to direct control of thought and action. (Block 1995, 381-382) Thus, one may be in either state, or both, and still be said to be having conscious experience.

Then there are the problem-cases in which there is no clear intuition as to whether a person is having conscious experience or not. One such case is that of blindsight. People with particular kinds of brain-damage report that they cannot see anything in one hemisphere of their visual field. They sincerely believe that they see nothing but darkness. Yet they are able to make guesses as to what appears in the half of their visual field which is supposedly dark with much better than chance accuracy. Thus, there is some evidence that they are "experiencing" something in the dark half, even if they are not aware of what it is. Another kind of problem-case is that of

individuals who have undergone brain surgery in order to alleviate unbearable amounts of pain localized to particular parts of their body. Afterwards, the patients report that they still know that they have the pain where the pain was originally. They still experience the pain in some way that they cannot really describe. They don't *feel* the pain, however. There is no immediate, non-inferential quality to the pain after the operation. There is some quality, however; otherwise, they would simply report the total absence of the pain post-surgery. And they seem to be experiencing this quality, even though they are not feeling the pain. These are just two of several kinds of cases in which it's difficult to decide whether experience is conscious or not.

So the question of which experiences to refer to as conscious is an open one. Some believe that we should not focus on the problem-cases, opting instead to attend to the cases which we pretheoretically refer to as "conscious experience." This position would state that, since we all know what kinds of experience should be referred to as "conscious," there really is no problem. Let's try to explain *them*, rather than the cases on which we cannot agree. But then there is the opposing position which simply denies that we have a good pretheoretical notion of "conscious experience." Thus, no headway is made at all.

The second prevalent disagreement probably derives from the first. If we cannot answer the question, "Which of our experiences properly fall under the term 'conscious experience?'" then we likely will not be able to answer the question, "What does the term 'consciousness' mean?" There are several disciplines working independently to resolve the problems associated with consciousness. Within each discipline, there are several distinct approaches to the problems. But, in fact, there is no agreed-upon definition of the thing these individuals are supposed to be researching. Flanagan writes, "There is no single orthodox concept of consciousness. Currently

afloat in intellectual space are several different conceptions of consciousness, many of them largely inchoate. … The picture of consciousness as a unified faculty has no special linguistic privilege, and none of the meanings of either ‘conscious’ or ‘consciousness’ wear any metaphysical commitment to immaterialism on its sleeve. The concept of consciousness is neither unitary nor well regimented, at least not yet” (Flanagan 1992, 13).

According to Crick, "Everyone has a rough idea of what is meant by consciousness. It is better to avoid a *precise* definition of consciousness because of the dangers of premature definition. Until the problem is understood much better, any attempt at a formal definition is likely to be either misleading, or overly restrictive, or both" (Crick 1994, 20). But we have already suggested that perhaps everyone's rough idea of what is meant by consciousness is a bit too rough. When we try to make the distinctions that science must make in order to investigate, the pretheoretical conception of consciousness may have to go.

Moreover, Güzeldere notes that there is no *one* problem of consciousness. He writes that the trouble may lie in "the fact that what passes under the term *consciousness* as an all too familiar, single, unified notion may be a tangled amalgam of several different concepts, each inflicted with its own separate problems" (Güzeldere 2002, 7). Furthermore, since consciousness is being investigated more and more in an interdisciplinary fashion, it is crucial that better characterizations of consciousness be arrived at. "After all, if we hope that anything toward a better understanding of consciousness will come out of the joint efforts of different disciplines, it is of utmost importance to minimize crosstalk and make sure that common terms actually point to the same referents. As a result, it seems even more imperative to look for and try to delineate the specific *conceptions* and *aspects* of consciousness under which the different problems arise"

(Güzeldere 2002, 8).

Despite the fact that Crick maintains that a formal definition would be premature—and we can agree with him on this point—it does not seem warranted to go on and uphold the pretheoretical notions of consciousness as adequate. As Güzeldere points out, we must disambiguate some of the tangled concepts associated with conscious experience. Perhaps an attempt at defining "consciousness" would be too ambitious, but a description of the concept seems in order to facilitate future relevant research.

I intend to provide a negative description—saying what “consciousness” *is not*—in order to eliminate the positions which are not useful at all for scientific research. Thus, I will have little to say about which of our experiences are conscious, what qualifies as falling under the concept “consciousness.” My goal is to move us closer to thinking of “consciousness” in a different way from what we have before.

Consider the breadth of experiences that we typically describe as conscious experience. There are mundane experiences—looking at a stoplight, hearing a co-worker talk about inflation, tasting the glue on the envelope flap. There are slightly more interesting experiences—hearing a child babble, feeling the smoothness of a polished stone, tasting baklava. Then there are the truly unique and enriching experiences—hearing a Beethoven sonata, tasting a perfectly aged Camembert, watching the sun set over the Grand Canyon. All three groups contain experiences, but there doesn’t seem to be much similarity between members of one group and another. In fact, there is a rather obvious difference—of degree. Presumably, one has a fuller experience of the child’s cooing than one does of the co-worker’s rant. One has a fuller experience of the sun setting than one does of the taste of envelope glue. So it seems rather questionable to group all of

these experiences—which differ greatly in degree—together under one term, "consciousness."

Not only is there a difference in degree, there is a difference in kind between various experiences. Consider the sensory modalities. We say that seeing a hamburger is a conscious experience in the same breath as we say that tasting a hamburger is a conscious experience. But there is a great difference between seeing and tasting. In fact, if we really try to compare the two modalities, it seems impossible to draw any correlations.[13] Seeing is completely different from tasting. And even if one sense (e.g. sight) seems to be more closely related to another (e.g. touch) than it does to a third (e.g. taste), this does not indicate that there are any similarities at all between the two seemingly more closely related senses. Thus, there is a difference in kind between the sensory modalities.

What about within a particular sensory modality? Is there a difference in kind between hearing a car horn and hearing a horn blast a resonant note in a symphony? If there is only a difference in degree, then there would just be some element in the hearing the car horn that was either more or less present in the hearing the symphony's horn. Is this the case?

I don't have the space needed to fully deal with the latter question. I will only say that there seems to me to be little similarity between the experience of hearing a car horn and the experience of hearing a note blasted in the middle of a symphony. Of course, some may be not agree with me, and claim that they perceive very little difference between the two. This could be because I am wrong or simply because those who disagree simply aren't perceiving a real difference that is present.

---

[13] I am ignoring those instances in which metaphors seem to indicate an overlap of modalities (i.e., tasting green or smelling purple.) Also, I am disregarding reports of synesthesia, for these kinds of experiences are outside the scope of this paper.

So, there are differences between the cases that we pretheoretically consider conscious and the marginal cases. There are differences in degree between certain mundane experiences, slightly less mundane experiences, and truly unique, edifying experiences. Moreover, there are differences in kind between the sensory modalities. And it seems plausible to suggest that there are some differences in kind even within the individual sensory modalities. With all these differences between various events we describe as experience, why do we feel tempted to refer to them all as “conscious experience?”

I don’t believe that "consciousness" is a truly unifying term. That is, I believe that “consciousness” does not pick out a common property that all experience possesses. It refers to the set of phenomena I have just described as wildly heterogeneous. So when we use the word "consciousness," we are not committing ourselves to a position in which there is a common characteristic which pulls all conscious experience together. It is a *façon de parler*, a device that enables us to speak about diverse phenomena, without enumerating them one by one. Since there is no "common thread" woven throughout all conscious experience, there is no reason to think that we can discuss "consciousness" in any coherent way, without discussing the fact that there is no unifying characteristic.

I suggest, therefore, that we think of "consciousness" as a family-resemblance term. Since there is no common characteristic belonging to each and every conscious experience, we should not use the term "consciousness" as though it referred to a collection of experiences marked by a unifying characteristic. This is not to say that *nothing* is referred to by "consciousness." We refer to the various experiences that we suspect are conscious. Science will bear out our suspicion or it will not. But, as a consequence of my *deflationist* description—

deflationist because it is an attempt to remove some of the *unnecessary* mysteriousness surrounding consciousness—science will not be investigating one natural kind … consciousness. It will be investigating several kinds, each with its own particular set of problems, as Güzeldere suggests should be the case.

So, to recap: there are intuitions regarding what experiences are conscious, but some have called into question whether they are reliable. Thus, intuitions get us moving, but they may misguide us at a certain point. We have to exercise some constraint. This is why I provide a deflationist picture of consciousness, as a negative description. *Consciousness is not homogeneous.* We have noted just how heterogeneous the collection we refer to as "consciousness" is. There is no common thread running through all conscious experience. To suppose there is would be to limit our description of conscious experience too much. As Crick notes, formal definitions at this point would likely be misleading or overly constrictive. I hope that the sketch I have provided is not misleading. I contend that the conception of consciousness that most philosophers trade on is indeed misleading, because consciousness is seen to be some uniform substance or property which brings together various experiences, so that they all share something in common—namely, being conscious experience. But this is circular and no help. I am rather confident that my sketch is not overly restrictive. Some may even accuse me of providing such a vague description that it does us no good, gets us no closer to a working definition. But I have not tried to provide a working definition. I have merely tried to describe the concept in a different way from the more common attempts.

Now, I have been arguing that my description of "consciousness" would be helpful in coming to an agreement on what we mean by "consciousness." Further, my conception seems to

fit in quite nicely with science's investigation of conscious experience. Neuroscience does not just study consciousness, simpliciter. It takes experience kind by kind. That is, it studies auditory phenomena, visual phenomena, gustatory phenomena, tactile phenomena, olfactory phenomena. It even studies and compares different experiences within the sensory modalities. It studies cognition, memory, and other mental processes in relation to phenomenal experience. Science proceeds piecemeal. My description of the concept "consciousness" seems more amenable to the scientific method of investigation than does a description that tries to lump all experience together via some particular common property.

What connection is there between my deflationist description and the zombie-argument? First, I would argue that Chalmers seems to be committed to consciousness being something uniform. He asks us to imagine that there is no consciousness in zombie world *Z*, without giving us much in the line of what he means by "consciousness." Thus, we don't know which position described above, if any, he subscribes to. If my suggestion is plausible, then the zombie-argument may be considered in light of the deflationist description.

So, let's assume that the deflationist description is a good one—no common characteristic may be articulated of conscious experience. Chalmers asks us to perform the zombie thought-experiment. We then come to the part where we are to stipulate that there is no consciousness in world *Z*. Now, however, we are at a loss as to what to do. We have no criterion for deciding what experience we are to stipulate out of existence. Chalmers asks us to imagine *all* conscious experience as being absent. This assumes that we have some way of deciding which experiences are conscious and which are not. Earlier I suggested that science will provide a lot of answers to this question, but it hasn't yet. Moreover, Chalmers will not accept such a reliance on science,

because he believes that science needs to adopt new laws in order to explain consciousness. We aren't imagining a science with psycho-physical laws here, however. Thus, he is not going to want to wait on science to help in determining what kind of experience we are stipulating the absence of in world *Z*.

Chalmers asks us to propositionally imagine *p's* being true, where *p* is the statement, "In world *Z* there is no conscious experience." But in order for me to be able to imagine *p*'s being true, I have to understand it, I have to know what it means to say that there is no conscious experience in *Z*. Does the phrase "conscious experience" contain unaware experience, like the collar's rubbing on the back of the neck? Or informationally rich, but phenomenally poor experience, like blindsight? But why would it include these kinds of experience? If there is no unifying characteristic common to these kinds of experience and, say, phenomenal experience, then why would we stipulate that they do not exist? It seems that Chalmers owes us an explanation of why we should imagine these kinds of experience as being absent from *Z,* and not other kinds.

Maybe he would just respond by saying, "All right. I am not asking you to imagine the absence of unaware experience or blindsight. I am only asking you to imagine the absence of phenomenal experience, the kind that we seem to all agree is conscious experience." But we have seen that even in relation to the experience we pretheoretically describe as conscious, there are wide differences. Why would we lump together smelling a rose with tasting envelope glue? Why does he include hearing a sleeve ruffle in the same class as hearing Mozart's Requiem? It seems he still owes us an explanation as to why it's *these* experiences that we should all agree are conscious.

But maybe he doesn't need any more than one experience. Perhaps he could just say, "Fine. I will not ask you to imagine a world bereft of several kinds or degrees of experience. I will only ask you to imagine a world physically identical to the actual world in which a token visual experience, which occurs in the actual world, is absent. It only takes one instance in which the physical does not entail the phenomenal for my conclusion to stand."

So, let's try this limited version of the zombie-argument. Imagine the possible world in which everything is physically the same as the actual world, but the visual experience I have of the screen before me is absent. My zombie-twin does not have the visual experience that I now have. Does this establish that materialism is false?

Actually, this seems like a straightforward scientific question. Can this token event of seeing the screen be fully explained in physical terms? Is the biological process that occurs when I am confronted with the computer screen sufficient to produce the phenomenal event of my seeing the screen? Why would scientific evidence need to be weighed against the finding of Chalmers's now-limited zombie-argument? The more limited the range of experience Chalmers is forced choose from, the more likely it is that science will be able to provide some answers to the question of whether it's possible for the physical processes to occur without the phenomenal experience. This limited new argument that Chalmers would have to adopt is not going to have the far-reaching implications he intends his original argument to have.

Thus, the zombie-argument gets no grip if we use the deflationist description of consciousness. This in itself I count as a strength. But I have given other reasons for supposing that there is no characteristic common to all conscious experience, independent of the zombie-argument. Thus, if one were to find my independent reasons good ones, the consequence of the

zombie-argument's losing its clutch on us would just be a bonus.

As I have made clear above, I believe that the zombie-argument, at its core, is not at all useful. I do think, however, that my negative description of consciousness is useful. I offer it up not because I don't find consciousness mysterious and elusive, but precisely because I do. Almost anyone who seriously considers conscious experience does, as well. It is incredibly mysterious, almost to the point of inexplicability. *Almost*. As I noted at the beginning of the paper, my position is not mysterian. I believe consciousness will be explained. But it will be explained piecemeal, each kind of experience studied separately, because many kinds will have to be explained in different ways. To think that consciousness as a whole will be explained via a cohesive, overarching theory is to engage in oversimplification. How then should consciousness be studied?

Conclusion

I believe that the correct method for studying phenomenal experience is something similar to what Owen Flanagan describes in *Consciousness Reconsidered.* There he writes

> I propose that we try the most natural strategy, what I call the *natural method,* to see if it can be made to work. Tactically, what I have in mind is this. Start by treating three different lines of analysis with equal respect. Give phenomenology its due. Listen carefully to what individuals have to say about how things seem. Also, let the

> psychologists and cognitive scientists have their say. Listen carefully to their descriptions about how mental life works and what jobs consciousness has, if any, in its overall economy. Finally, listen carefully to what the neuroscientists say about how conscious mental events of different sorts are realized, and examine the fit between their stories and the phenomenological and psychological stories. (Flanagan 1992, 11)

Notice that Flanagan does not advocate a complete disregard for phenomenology. The neuroscientists may explain every detail of the brain and still be susceptible to charges that the explanatory gap is still there, and as wide as ever. That is why we must include the "what it's like" element in research done on consciousness. Similarly, we should not eliminate the investigation of what function certain cognitive capacities have. If we were to do that, we might be missing a vital piece of the puzzle. The natural method, as Flanagan calls it, would require each of the three lines. Ignoring any one of them would be to our peril.

Also notice that Flanagan does not include thought-experiments in his vision of a good, useful method for studying consciousness.

Where does that leave the zombie-argument? Recall that if it's correct, then science as it stands will need to be overhauled, so that it can make room for non-physical entities in the world. The natural method does not seem to be taking into account any sort of overhaul of science. Indeed, it seems to be committed to a physicalist picture. So what's to be done?

According to John Searle, thought-experiments like the zombie-argument are counterproductive and even pernicious. In order to answer arguments for dualism, one must accept the vocabulary of the dualist, which poisons the discourse from the start. "In denying the

dualist's claim that there are two kinds of substances in the world or denying the property dualist's claim that there are two kinds of properties in the world, materialism inadvertently accepts the categories and the vocabulary of dualism" (Searle 1994, 54). So, rather than getting on with the hard work of pursuing the explanations of particular phenomenal experience, materialists have to defend against attacks like Chalmers's zombie-argument. But, in the process, they have to assume the kind of language that will not be helpful in the relevant research. And so, they consume a lot of time and energy which would be better spent in other areas.

I believe Searle is right here. These kinds of arguments have no place, or at least no significant place, in philosophy of mind. I think matters of ontology, reflections on what kinds of things there are in the world, have a place in metaphysics, but in the philosophy of mind, I think it's a serious neurosis we should be cured of, once and for all.

Bibliography

Balog, K. (1999) "Conceivability, Possibility, and the Mind-Body Problem," *The Philosophical Review*, 108, 4: 497-528.

Block, N. (2002) "On a Confusion about a Function of Consciousness," in Block, N., Flanagan, O., and Güzeldere, G. (eds.), *The Nature of Consciousness*. Cambridge, MA: MIT Press.

Bonjour, L. (1998) *In Defense of Pure Reason*. New York: Cambridge University Press.

Chalmers, D. J. (1996) *The Conscious Mind.* New York: Oxford University Press.

Chalmers, D. J. (2002). "Does Conceivability Entail Possibility?" in Gendler, T. S., and Hawthorne, J. (eds.), *Conceivability and Possibility*. Oxford: Clarendon Press.

Cottrell, A. (1999) "Sniffing the camembert: On the conceivability of zombies," *Journal of Consciousness Studies* 6, 1: 4-12.

Crick, Francis. (1994) *The Astonishing Hypothesis.* New York: Charles Scribner's Sons.

Dennett, D. C. (1991) *Consciousness Explained.* Boston, MA: Back Bay Books.

Dennett, D. C. (1995) "The Unimagined Preposterousness of Zombies," *Journal of Consciousness Studies* 2, 4: 322-326.

Dietrich, E., and Gillies, A. (2001) "Consciousness and the Limits of our Imaginations," *Synthese* 126: 361-381.

Drestke, F. (1997) *Naturalizing the Mind.* Cambridge, MA: MIT Press.

Flanagan, O. (1992) *Consciousness Reconsidered.* Cambridge, MA: MIT Press.

Güzeldere, G. (2002) “The Many Faces of Consciousness: A Field Guide,” in Block, N., Flanagan, O., and Güzeldere, G. (eds.), *The Nature of Consciousness*. Cambridge, MA: MIT Press.

Hill, C. S., and McLaughlin, B. P. (1999) “There Are Fewer Things in Reality Than Are Dreamt of in Chalmers’s Philosophy,” *Philosophy and Phenomenological Research* LIX, 2: 445-454.

Kirk, R. (1974) “Sentience and Behaviour,” *Mind* 83, 329: 43-60.

Levine, J. (1998) “Conceivability and the Metaphysics of Mind,” *Nous* 32, 4: 449-480.

Loux, M. (1979) *The Possible and the Actual*. Ithaca, NY: Cornell University Press.

Marcus, E. (2004) “Why Zombies are Inconceivable,” *Australasian Journal of Philosophy* 82, 3:477-490.

Marton, P. (1997) “Zombies versus Materialists: The Battle for Conceivability,” *Southwest Philosophy Review* 14: 131-138.

McGinn, C. (1989) “Can We Solve the Mind-Body Problem?” *Mind* 98: 349-366.

Perry, J. (2001) *Knowledge, Possibility, and Consciousness*. Cambridge, MA: MIT Press.

Seager, W. (unpublished) "Are Zombies Logically Possible? And Why It Matters.”

Searle, J. (1994) *The Rediscovery of the Mind.* Cambridge, MA: MIT Press.

Shoemaker, S. (1999) “On David Chalmers’s ‘The Conscious Mind,’” *Philosophy and Phenomenological Research* 59, 2: 439-444.

Wittgenstein, L. (1958) *Philosophical Investigations.* Englewood Cliffs, NJ: Prentice Hall.

Yablo, S. (1993) “Is Conceivability a Guide to Possibility?” *Philosophy and Phenomenological Research* 53, 1: 1-42.