# What Was That Like? Intuitions and the Epistemology of Consciousness

Abstract

I argue that physicalists have been too conciliatory in granting that certain classic thought experiments about consciousness such as Mary the color scientist, color spectrum inversion, and zombies provide strong prima facie support for epiphenomenal anti-physicalism. While these thought experiments may suggest that we are intuitive epiphenomenal anti-physicalists when taken individually, when they are appropriately combined, they suggest that epiphenomenal anti-physicalism leads to a version of phenomenal skepticism according to which (i) we cannot know how our states of phenomenal consciousness compare and contrast and (ii) we cannot know how our first-order beliefs about our states of consciousness compare and contrast. Insofar as comparative phenomenal skepticism is a deeply counter-intuitive view, our intuitions about consciousness are far more equivocal than they are widely thought to be. There simply may be no one metaphysical view that should qualify as their obvious champion.

Key words: consciousness, epistemology, intuitions, physicalism, anti-physicalism

By: Brandon Ashby

# 1. Introduction

 Epiphenomenal anti-physicalism claims that experiences (or phenomenal properties/qualia) are non-physical and that there is no phenomenal-to-physical causation. Even among physicalists about consciousness, it is widely accepted that certain classic thought experiments such as Mary the color scientist, color spectrum inversion, and zombies provide strong prima facie support for epiphenomenal anti-physicalism.

In this paper, I shall argue that physicalists about consciousness have been too conciliatory in granting that the classic thought experiments provide unqualified prima facie support for epiphenomenal anti-physicalism. I shall not argue that the thought experiments are inconceivable or unimaginable (see Dennett, 1995, 2005, 2007; Kirk, 1999). Rather, my claim is that the classic thought experiments may suggest that we are intuitive epiphenomenal anti-physicalists when they are taken individually, but, when they are appropriately combined with one another, they also suggest that epiphenomenal anti-physicalism leads to a version of phenomenal skepticism according to which (i) we cannot know how our states of phenomenal consciousness compare and contrast with one another, and (ii) we cannot know how our first-order beliefs about our states of consciousness compare and contrast with one another. Insofar as *comparative phenomenal skepticism* is a deeply counter-intuitive view, our intuitions about consciousness are far more equivocal than they are widely thought to be, and there simply may be no one metaphysical view that can qualify as their obvious champion. To make my case, I shall focus on the intersection of two key debates in the philosophy of consciousness: the paradox of phenomenal judgment and the nature of phenomenal concepts.

 A number of physicalists argue that epiphenomenal anti-physicalism entails *phenomenal skepticism simpliciter*: we cannot know what, if anything, it is like to be us or what it has been like to be us (Campbell, 2012, 2003; Dennett, 1991, 398-406; Jackson, 1995; 2004; Shoemaker, 1975; Watkins, 1989). As we shall see in §2, *the paradox of phenomenal judgment*, as this argument is known, is extremely plausible so long as we presuppose:

*Content physicalism*: belief contents supervene on the physical/functional properties of subjects (and their environments) with conceptual necessity.

The problem faced by the paradox of phenomenal judgment is that content physicalism is at odds with one of the most popular views on the nature of phenomenal concepts, one that, as we shall see, receives powerful motivation from the classic thought experiments (see §3). Following David Chalmers, I shall call it *the theory of pure phenomenal concepts*. The theory is widely accepted by both physicalists and anti-physicalists alike (Balog, 2012; Gertler, 2007; Horgan & Kriegel, 2007; Levine, 2001; Papineau, 2002; Shoemaker, 1996; Block, 2006). And it has been put to use in a range of debates beyond those over the epistemology of epiphenomenal anti-physicalism (Barz, 2017; Silins, 2014; Ellis, 2010).

While the theory of pure phenomenal concepts has many defenders, here I shall focus on David Chalmers’ presentation because he appeals to an intuitive contrast between Mary the color scientist and her color spectrum inverted counterpart to motivate the theory and to illustrate the precise nature of the problem that it creates for the paradox of phenomenal judgement (BonJour, 2003; Butler, 2011; Duncan, 2017; and Nagasawa, 2010 offer arguments that carry similar consequences for the paradox but do not appeal to the Mary/Inverted Mary contrast).[[1]](#footnote-1) This make the connections between the classic thought experiments, the paradox of phenomenal judgment, and nature of phenomenal concepts particularly clear.

While the contrast between Mary and Inverted Mary undermines content physicalism (§3), I shall argue that it actually supports a nearby position:

*Content-structure physicalism*: the overall *form* or *structure* of belief contents supervene on the physical/functional properties of subjects (and their environments) with conceptual necessity.[[2]](#footnote-2)

And I shall argue that we can construct a version of the paradox of phenomenal judgment that only requires content-structure physicalism, not content physicalism. I call it *the quarantine problem*. It claims that epiphenomenal anti-physicalism entails comparative phenomenal skepticism: we cannot know how our experiences compare and contrast with one another and we cannot know how our first-order beliefs about our states of consciousness compare and contrast (see §4). The upshot is that the very same thought experiments that are widely thought to offer such strong prima facie support for epiphenomenal anti-physicalism and undermine the paradox of phenomenal judgment also show that epiphenomenal anti-physicalism faces a problem with comparative phenomenal skepticism. Ultimately, our intuitions fail to tell a neat and tidy story.[[3]](#footnote-3)

 I begin by presenting the paradox of phenomenal judgment in §2. I then introduce the Mary/Inverted Mary contrast to show how it motivates the theory of pure phenomenal concepts and undermines the paradox of phenomenal judgment (§3). I introduce the quarantine problem in §4, respond to objections in §5, and draw more general conclusions in §6.

# 2. The Paradox of Phenomenal Judgment

While there are subtle differences in presentation, Neil Campbell (2003, 2012), Daniel Dennett (1991, 398-406), Frank Jackson (1995, 2004), Sydney Shoemaker (1975), and Michael Watkins (1989) all offer a common argument against epiphenomenal anti-physicalism. Jackson’s version of the argument appeals to *the causal closure of the physical domain*—the principle that fundamental physics can provide sufficient explanations for physical events insofar as there are any explanations of those events to be had:

[W]e now know enough to know that were there anything nonphysical about our psychology, it would be screened out by our physical natures in the sense that *no indicative traces of it would survive in memory, in reports, in articles called ‘Epiphenomenal Qualia’ and so on*. This creates an insurmountable epistemic problem for the views of my former [epiphenomenal anti-physicalist] self. (2004, xvi; emphasis added)

Dennett’s version concerns the evidential import of our claims about consciousness:

Suppose…that Otto insists that he (for one) has epiphenomenal qualia. Why does he say this? Not because they have some effect on him… By the very definition of epiphenomena… Otto’s heartfelt avowals that he has epiphenomena could not be evidence for himself or anyone else that he does have them, *since he would be saying exactly the same thing even if he didn’t have them*. (1991, 402-3; emphasis added)

And Watkins focuses directly on justification:

If qualia are not causally efficacious, thenmy beliefs and memories *would be just as they are whether there were qualia or not.* Beliefs about qualia cannot be justified on the basis of qualitative experiences *since those experiences do not cause those beliefs…* (1989, 160; emphasis added)

Despite the differences in presentation, we can regiment the argument as follows:

1. If epiphenomenal anti-physicalism is true, then the contents of our beliefs about our experiences do not depend on what experiences, if any, we have.

For example, even if we had all been zombies, we still would have believed that there is something that it is like to taste chocolate.

1. If the contents of our experiential beliefs do not depend on what experiences, if any, we have, then they are unjustified.
2. Without justification, beliefs cannot qualify as knowledge.
3. Hence, epiphenomenal anti-physicalism entails phenomenal skepticism.
4. Phenomenal skepticism is absurd.
5. Hence, epiphenomenal anti-physicalism is false.

While (1) may or may not seem compelling on its own, it receives considerable support from:

*Content physicalism*: belief contents supervene on the physical/functional properties of subjects (and their environments) with conceptual necessity.

If content physicalism is correct, then non-physical properties could only influence what we believe by influencing the physical/functional bases of our beliefs. But epiphenomenal anti-physicalism denies that experiences can have physical effects. So, experiential belief content gets *screened off* from our experiences by its physical/functional nature. Hence, (1).

As we shall see in §3, there are compelling reasons to reject content physicalism and (1) so long as we grant the conceivability of the classic thought experiments. Chalmers’ contrast between Mary and Inverted Mary gives us strong reasons to embrace the theory of pure phenomenal concepts, which, as we shall see, undermines content physicalism. In turn, I shall argue that there is a modified version of (1) that only presupposes content-structure physicalism, not content physicalism:

(1\*) If epiphenomenal anti-physicalism is true, then our beliefs about how our experiences compare and contrast do not depend on whether or not the compared or contrasted experiences actually are similar to or different from one another.

By appropriately modifying (2)-(6), one can argue against epiphenomenal anti-physicalism on the grounds that it entails comparative phenomenal skepticism.

# 3. Pure Phenomenal Concepts and The Paradox of Phenomenal Judgment

## 3.1. Pure Phenomenal Concepts

Chalmers contrasts Mary the color scientist with her color spectrum inverted twin. Both Marys are physical/functional duplicates of each other living in physically/functionally equivalent worlds. And they are raised without any exposure to color until, eventually, each is given a ripe tomato. Mary then has the sort of color experience that we have when looking at red tomatoes, and Inverted Mary has the sort of color experience that we have when looking at green cucumbers.

Intuitively, each Mary learns something new. But they do not learn the same thing; their newly acquired beliefs about what it is like for them to see red concern different kinds of experiences. Nevertheless, Chalmers claims, we have no difficulty imagining that the two Marys remain physical/functional duplicates of one another throughout their lives. So, we cannot be thinking of their divergent experiences as making a *causal difference* to their physical/functional constitutions, since we do not imagine any physical/functional differences between them. Hence, content physicalism must be false—belief content does not conceptually supervene on physical/functional properties—and so we need an account of how the contents of our experiential beliefs could depend on our experiences in the absence of any accompanying physical/functional differences. Chalmers develops his theory of pure phenomenal concepts to meet this need, and, while different authors have different motivations for doing so, the sort of theory that Chalmers advocates has proven to be extremely popular (Balog, 2012; Gertler, 2007; Horgan & Kriegel, 2007; Levine, 2001; Papineau, 2002; Shoemaker, 1996; Block, 2006). Nevertheless, I shall focus on Chalmers’ presentation because his appeal to the contrast between Mary and Inverted Mary makes the role that the classic thought experiments have to play in the current debate especially clear.

According to Chalmers, there are two kinds of pure phenomenal concepts: direct and standing. *Direct phenomenal concepts* contain a “slot” in their contents that token experiences—such as Mary’s first experience of red—can “fill in”. A direct phenomenal concept then picks out either that particular experience or the type of experience that the token experience is an instance of. In Chalmers’ words: “[A] phenomenal quality can be thought of as filling a slot that is left open in the content of a direct phenomenal concept and thus contributing its content” (Chalmers, 2010: 294). So, Mary and Inverted Mary each coin a direct phenomenal concept when looking at their tomatoes, and each thinks a thought of the form, ‘*Direct  is what it is like to see red objects*’, where the ‘**’ represents the slot in the contents of their direct phenomenal concepts. Because different kinds of experiences fill in the slots in Mary’s and Inverted Mary’s direct phenomenal concepts, their thoughts differ in content despite their physical/functional equivalence.

Direct phenomenal concepts can only exist as long as the token experiences that fill in the slots in their contents do. When a direct phenomenal concept ceases to exist because its associated experience token ceases to exist, it will typically be succeeded by a *standing phenomenal concept* (Chalmers, 2010: 271). Chalmers claims that standing phenomenal concepts inherit their semantic properties from whatever direct phenomenal concept they succeed (2010: 271-2).[[4]](#footnote-4) This is to explain how Mary and Inverted Mary can continue to have distinct thoughts from one another about their color experiences even when they are not currently having any color experiences.

Some readers may question Chalmers’ use of Inverted Mary on the grounds that, if one claims that the phenomenal domain supervenes on the physical/functional domain with nomological necessity, then Inverted Mary’s scenario is counternomic. For the purposes of this paper, I want to set any such objections aside. After all, Campbell, Dennett, Jackson, Shoemaker, and Watkins all discuss zombies and other scenarios that will be counternomic if the phenomenal supervenes on the physical/functional with nomic necessity. So, I shall grant Chalmers and his interlocutors that potentially counternomic scenarios can matter for the epistemology of consciousness, since this assumption is shared by both sides of the debate.

## 3.2. Chalmers’ Response to The Paradox of Phenomenal Judgment[[5]](#footnote-5)

The first premise of the paradox of phenomenal judgment reads:

1. If epiphenomenal anti-physicalism is true, then the contents of our beliefs about our experiences do not depend on what experiences, if any, we have.

Chalmers argues that (1) is false by challenging content physicalism. Mary’s and Inverted Mary’s direct phenomenal concepts have different contents because Mary’s concepts have token red experiences filling in their slots wherever Inverted Mary’s concepts have token green experiences filling in their slots, and vice versa. Since standing phenomenal concepts inherit their contents from the direct phenomenal concepts they succeed, the two Marys’ standing phenomenal concepts differ in content as well. Consequently, the contents of the two Marys’ beliefs that involve pure phenomenal concepts will depend on what sorts of experiences they have. Since the two Marys are physical/functional duplicates of one another living in physically/functionally equivalent worlds, their experiences cannot be making a causal difference to the physical/functional basis of what they think. Hence, content physicalism and (1) are false.

# 4. Periodically Inverting Mary and The Quarantine Problem

Chalmers’ contrast between Mary and Inverted Mary reveals important and compelling intuitions, but I shall argue that those intuitions run afoul of the purposes that Chalmers puts them to. While the Mary/Inverted Mary contrast undermines content physicalism, as we shall see, it actually supports content-structure physicalism, the claim that the overall form or structure of belief contents conceptually supervene on physical/functional properties, which in turn supports (1\*). So, the very same resources that Chalmers marshals against the paradox of phenomenal judgment can be used to motivate a revised version of it.

The basic problem that epiphenomenal anti-physicalists face is that the differences in what we would have thought if our experiences had differed (without any accompanying physical/functional differences that could allow the overall structure of our thoughts to differ) will be limited to our pure phenomenal concepts alone. The rest of our trains of thought will go unchanged. This quarantining of possible differences in content to pure phenomenal concepts prevents our experiential beliefs about the similarities and differences between our experiences from depending on whether or not those experiences actually are similar to or different from one another. While my arguments in this section will focus on experiences had at different times, in §5.1 I show how to generalize the argument to experiences had at the same time.

## 4.1. Arts and Crafts and Concepts

I want to begin by offering a metaphor for understanding the quarantine problem before offering some examples of it (§§4.2-4.4) and arguing that it only applies to epiphenomenal anti-physicalism (§4.5). Imagine that we write out the contents of a subject’s thoughts on a piece of paper. To represent direct phenomenal concepts, we write ‘direct’ and punch a hole nearby. The hole represents the slot in the direct phenomenal concept’s contents. For standing phenomenal concepts, we write ‘standing’ and draw an arrow back to whatever hole is associated with the direct phenomenal concept that the standing phenomenal concept succeeded. For instance, the paper representing Mary’s and Inverted Mary’s post-color thoughts looks like this:

Figure



By laying the paper over differently painted canvases, we can represent how the subject’s experiential beliefs depend on her experiences. One canvas represents the subject’s actual color experiences, while the others represent ways that the subject’s color experiences could have gone but did not, for instance, if she had been a color spectrum invert if she is not, or vice versa.[[6]](#footnote-6) Whatever color of paint shows through a hole in the paper represents the sort of color experience that the associated direct phenomenal concept does or would refer to if the subject were to have the sort of color experience represented by that color of paint. Whatever color of paint an arrow points to will represent the sort of color experience that the associated standing phenomenal concept does or would refer to if the subject were to have the sort of color experience represented by that color of paint.

Figure

The paper and canvas combinations representing Mary’s (left) and Inverted Mary’s (right) thoughts.

 

The words written on the paper, where the holes are punched, and how the arrows are drawn all represent the ways in which Mary’s and Inverted Mary’s thoughts have the same structure or form. After all, when we imagine Mary and Inverted Mary, we imagine that they both end up thinking structurally equivalent thoughts that only differ in what sorts of experiences those thoughts are about. For instance, we imagine that they will both think something like, ‘*Ah, so direct is what it is like to see red*’ when they first see a ripe tomato. Consequently, we can use a single piece of paper to represent the equivalent structure of their thoughts, and we can lay that paper over different canvases to capture the differences in what sorts of experiences their thoughts concern. More generally, what Mary and Inverted Mary suggest is that physical/functional duplicates will have structurally equivalent thoughts—that is, they support content-structure physicalism.[[7]](#footnote-7)

## 4.2. First-Order Phenomenal Comparisons

Let us consider Periodically Inverting Mary.[[8]](#footnote-8) She is a physical/functional duplicate of Mary and Inverted Mary and she lives in a world that is physically/functionally equivalent to their worlds, yet she alternates between being a normal subject and a color spectrum invert.[[9]](#footnote-9) We can note at the outset one problem that she faces: she does not know what it is like to see red because there is no *one way* that it is like for her to see red. Red objects induce phenomenally red experiences in her some of the time, and phenomenally green experiences the rest of the time. Setting this issue aside, there are a number of different kinds of judgement that are problematic for her. Each serves as an example of the quarantine problem.

Say that Mary looks at a tomato at time t1 and forms a direct phenomenal concept of her color experience. Then she closes her eyes and her direct phenomenal concept ceases to exist. At t2, she opens her eyes and forms another direct phenomenal concept. Then she uses her standing phenomenal concept of her experience at t1 and her direct phenomenal concept of her experience at t2 to think that her earlier and later experiences are the same:

*Figure 3*



Inverted Mary forms a structurally equivalent belief:

*Figure 4*



Periodically Inverting Mary also forms a structurally equivalent belief, but her color spectrum inverts between t1 and t2. So, she ends up believing that her earlier experience of phenomenal red was the same as her later experience of phenomenal green:

Figure



Her belief is false.

The problem just illustrated generalizes. Comparative and contrastive thought components like:

 …*is the same sort of experience as…*

 *…is a similar sort of experience to…*

 *…is a totally different sort of experience from…*

are not pure phenomenal concepts and are not built out of pure phenomenal concepts.[[10]](#footnote-10) So, if any one of the three Marys deploys one of these thought components, then the other two will as well *no matter what their experiences are like*. Comparative and contrastive thought components are structural features of subjects’ thoughts and so are invariant between physical/functional duplicates. In terms of the metaphor, they get printed on the page without any of the holes or arrows that could allow their contents to depend on the color pattern of the underlying canvas.

Even total thoughts that concern patterns of similarity and difference in our experiences, such as:

 *I am not a periodic color spectrum invert.*

form part of the structure of the three Marys’ thoughts. If Mary or Inverted Mary believe that they are not periodic color spectrum inverts, then Periodically Inverting Mary does too. And, if Periodically Inverting Mary believes that:

 *My color spectrum is periodically inverting*.

then Mary and Inverted Mary will as well. When it comes to comparative and contrastive experiential beliefs, Periodically Inverting Mary must be wrong if Mary and Inverted Mary are right, and vice versa.

Chalmers’ contrast between Mary and Inverted Mary shows that:

1. If epiphenomenal anti-physicalism is true, then the contents of our experiential beliefs do not depend on our experiences.

is too strong of a claim. However, (1\*) still stands:

(1\*) If epiphenomenal anti-physicalism is true, then our beliefs about how our experiences compare and contrast do not depend on whether or not the compared or contrasted experiences actually are similar to or different from one another.

This claim only requires that the overall *structure* or *form* of belief contents conceptually supervene on physical/functional properties. And, unlike content physicalism, the intuitive contrast between Mary and Inverted Mary supports content-structure physicalism (see §4.1). So, epiphenomenal anti-physicalists face the quarantine problem.

## 4.3. Meta-Cognition Without Pure Phenomenal Concepts

The quarantine problem also threatens our abilities to know how *our thoughts about our experiences* compare and contrast with one another. It threatens comparative phenomenal skepticism at the meta-cognitive level.

Say that Mary looks at the sky while going for a walk. She thinks:

Figure



Then she gets distracted by a brief phone call and forgets all about her color experience. Afterwards, her mind begins to wander, and she ends up thinking:

Figure



But then it hits her:

Figure



Mary’s meta-judgement is true. She has indeed thought about how pleasant one particular kind of experience is twice over in short order.

Periodically Inverting Mary’s color spectrum inverts when she takes the phone call. So her thought process looks like this:

Figure



Her meta-judgement is false. Given the strictures of epiphenomenal anti-physicalism, if Mary forms the meta-judgement, then Periodically Inverting Mary must as well. Despite their experiential differences, they remain physical/functional duplicates of one another. So their thoughts are structurally equivalent. In terms of the metaphor, the pages that represent their thoughts are equivalent; only the underlying canvases can differ. So, our second-order thoughts that our first-order experiential beliefs are similar to or different from one another do not depend on whether or not those first-order beliefs actually are similar to or different from one another. It is the quarantine problem at the meta-cognitive level.

## 4.4. Meta-Cognition with Pure Phenomenal Concepts

The meta-judgement contained no pure phenomenal concepts. But we can construct an example that does. Say that Mary looks at a tomato for a few moments and thinks:

Figure



Mary’s common content judgement is true. But Periodically Inverting Mary ends up thinking:

Figure 11



Periodically Inverting Mary’s common content judgement includes a pure phenomenal concept—so its content depends on her experiences—but she still ends up thinking that her last two thoughts concerned a common experience when they did not. Even when we use pure phenomenal concepts to compare and contrast our experiential beliefs, our higher-order thoughts that our first-order experiential beliefs are similar to or different from one another do not depend on whether or not those first-order experiential beliefs actually are similar to or different from one another. The quarantine problem is quite general.

## 4.5 The Quarantine Problem Only Affects Epiphenomenal Anti-Physicalists

Some problems are everyone’s problem. The quarantine problem is not one of them. If we were to accept anti-physicalism but reject epiphenomenalism—call this *interactionist anti-physicalism*—then the three Marys would not remain functional duplicates of one another in their post-color lives; their divergent color experiences would causally impact them in different ways. Consequently, their thoughts could differ structurally. In terms of the metaphor, there is no requirement that we use the same piece of paper, since they are not functional duplicates. Mary could think:

Figure



while Inverted Mary thinks:

Figure



and Periodically Inverting Mary thinks:

Figure



Whether or not the three Marys’ experiential beliefs will differ in structure depends on how their different experiences causally impact their physical/functional states. But interactionist anti-physicalists are under no special burden to rule out the possibility that, even though the three Marys’ experiences causally impact the physical/functional underpinnings of their psychologies, those experiences never do so in a way that produces structurally different experiential beliefs.

Moreover, if we were to accept physicalism, then the physical differences that serve as the categorical bases for their experiential differences could result in their having structurally different thoughts from one another. Of course, there is the possibility that these physical differences never suffice for a structural difference in their experiential beliefs. But physicalists are under no special burden to rule this out, since physical differences are precisely the sorts of things that philosophers think can suffice for differences in cognition.[[11]](#footnote-11)

Now consider epiphenomenal anti-physicalism. Because phenomenal properties are non-physical, a subject can have different experiences without differing in her physical/functional makeup. Because experiences do not causally impact the physical domain, then, if the subject’s experiences had differed, those differences would not have caused any physical/functional differences in the subject. So, on the epiphenomenal anti-physicalist view, if we had been periodic color spectrum inverts, we would not have differed physically/functionally from how we are.[[12]](#footnote-12)

As noted in §4.1, physical/functional duplicates have structurally equivalent thought processes. Consequently, there is no possibility that Periodically Inverting Mary’s experiential beliefs could differ structurally from those of Mary and Inverted Mary. Given the strictures of epiphenomenal anti-physicalism, we must use a single piece of paper laid over different canvases to represent how the three Marys’ thoughts would have differed if their experiences had differed. So, if we had been periodic color spectrum inverts, we still would have compared and contrasted our experiences *in* *exactly the same ways that we do*, and likewise for our comparative and contrastive thoughts about our first-order experiential beliefs. Pure phenomenal concepts do not show that epiphenomenal anti-physicalists can avoid comparative phenomenal skepticism.

# 5. Objections and Replies

I have argued that only epiphenomenal anti-physicalists suffer from the quarantine problem: the differences in what we would have thought if our experiences had differed are limited to the contents of our pure phenomenal concepts alone. The conceptual repertoire that we use to compare and contrast our experiences and our experiential beliefs, such as:

 *…was the same sort of experience as…*

are part of the psychological structure that is invariant across physical/functional duplicates. So, we get (1\*): our thoughts that our experiences are similar to or different from one another do not depend on whether or not our experiences actually are similar to or different from one another. And likewise for our higher-order thoughts about the similarities and differences between our first-order experiential beliefs. This allows us to reconstruct a version of the paradox of phenomenal judgment that does not rely on content physicalism, only content-structure physicalism, which is supported by the contrast between Mary and her color spectrum inverted counterparts. In the rest of this section, I respond to possible objections.

## 5.1. Feature Not a Bug

Epiphenomenal anti-physicalists could bite the bullet. To motivate this response, they could draw a parallel between the problem of other minds and comparative phenomenal skepticism for experiences had at different times, the problem of other times, as it were.

This is a large bullet to bite, however. The epiphenomenal anti-physicalist would be conceding that we can never know whether or not what it was like to be us a moment ago was even remotely similar to what it is like to be us here and now. Indeed, since we can always run the quarantine problem in terms of Periodically Zombified Mary, who alternates between being a normal subject and being a zombie, the epiphenomenal anti-physicalist would end up conceding that we cannot even know if we were conscious a moment ago.

 Moreover, we can construct a synchronic case of partial color spectrum inversion. Say that Mary enters a psychology study. The experimenters present her with pairs of paint chips. One chip always appears on her left, the other on her right. For each pair, Mary is asked to say whether or not her experiences of the two chips are equivalent. At various points throughout the experiment, she thinks to herself:

Figure

*My direct  experience of the left chip is the same as my direct  experience of the right chip*.

 Now consider *Hemifield Mary*. She is just like all the other Marys except that at all times her right visual field is color spectrum inverted and her left visual field is not. When Hemifield Mary participates in the study, she ends up thinking:

Figure

*My direct  experience of the left chip is the same as my direct  experience of the right chip*.

Here we have a healthy subject paying full attention to two very different color experiences had simultaneously, and yet she fails to notice the difference. The epiphenomenal anti-physicalist may be willing to accept comparative phenomenal skepticism for experiences had at different times, but comparative phenomenal skepticism for experiences had at the same time should be unacceptable on anyone’s accounting.

## 5.2. A Law of Nature

Chalmers’ use of pure phenomenal concepts is not his first attempt to answer the paradox of phenomenal judgement. In *The Conscious Mind*, Chalmers introduced his “dancing qualia” thought experiment to argue that our experiences supervene on our fine-grained functional states with nomic necessity (1996, 266-273). He calls this law of nature *the principle of organizational invariance*. [[13]](#footnote-13) On the resulting view, zombies, color spectrum inverts, and the like are metaphysically possible, but nomologically impossible. This principle of organizational invariance can be used to formulate a single response to both the paradox of phenomenal judgment and the quarantine problem.

The paradox of phenomenal judgment claims that we still would have believed that we are phenomenally conscious even if we had all been zombies. If there is a principle of organizational invariance, then any nomologically possible worlds in which we lack consciousness are worlds in which our functional states differ from what they are. So long as those functional differences result in appropriate differences to what we believe, then we will not believe that we are conscious in those worlds. The epiphenomenal anti-physicalist can then claim that nomologically possible worlds are what matter most in epistemology.

Similarly, the quarantine problem claims that Mary would have made the problematic sorts of judgements that Periodically Inverting Mary did if she had been a periodic color spectrum invert. If there is a principle of organizational invariance, then, in any nomologically possible world where Mary has alternating experiences of red and green when looking at ripe tomatoes, her functional states differ from what they actually are. So long as those functional differences result in appropriate differences in the structure of what she thinks, then she will not form the problematic beliefs that Periodically Inverting Mary did.

I have two responses to this *law-of-nature objection*. First, Chalmers himself is less than satisfied with his earlier work on the paradox of phenomenal judgement. He says, “I now think that discussion is at best suboptimal…” and that his theory of pure phenomenal concepts “can be viewed in part as a replacement for that discussion.” (2010: Ch. 8, fn. 7). And elsewhere he says that his use of pure phenomenal concepts, “provides a better response to the ‘paradox of phenomenal judgment’ than I gave in *The Conscious Mind*” (2010, xxi).

Since the law-of-nature objection makes the same basic move against both the paradox of phenomenal judgment and the quarantine problem, we should expect it to succeed against one if it succeeds against the other; the quarantine problem is, after all, a modified reconstruction of the paradox. So, if there was some residual threat posed by the paradox of phenomenal judgment that the principle of organizational invariance could not resolve, then the quarantine problem will resuscitate that threat. On the other hand, if the law-of-nature objection defeats the quarantine problem, then it also defeats the paradox of phenomenal judgment and so there was never any need for Chalmers to “replace” his “at best suboptimal” response in the first place.

Second, even if successful, the law-of-nature objection would not show us that pure phenomenal concepts allow the epiphenomenal anti-physicalism to provide a satisfactory epistemology of consciousness; rather, it would show us that *the principle of organizational invariance*—not pure phenomenal concepts—is what saves epiphenomenal anti-physicalism. After all, we did not even need to *mention* pure phenomenal concepts when offering the law-of-nature objection. And, as already noted, Chalmers’ assessment of his earlier work is less than sunny.

## 5.3. Panprotopsychism

Finally, consider *panprotopsychism*. According to panprotopsychism, there is a single ontological category from which both the physical domain and phenomenal consciousness coemerge as distinct manifestations. Think of how two TV screens showing the same sporting event from different angles would present distinct but non-independent images. So, it would be incorrect to think of differences in experience as *causing* physical differences, or vice versa. The two TV screens do not causally influence one another. So, panprotopsychism can be construed as an epiphenomenalist view. Moreover, depending upon exactly how the physical and phenomenal domains coemerge from the more fundamental layer of reality, it may be metaphysically impossible for our experiences to vary independently of our physical/functional states in the sorts of ways required for the quarantine problem to arise.

Here is the problem. If a version of panprotopsychism accepts the metaphysical possibility of color spectrum inverts and zombies, then it is an anti-physicalist position in spirit. Otherwise, it looks much closer to physicalism. And, if advocates of panprotopsychism grant the metaphysical possibility of full-blown color spectrum inverts and zombies, then it is hard to see how they can justify claiming that partial inverts and partial zombies (like Periodically Inverting Mary) are metaphysically impossible. So, either panprotopsychism appears near enough to physicalism as makes no difference—it denies the metaphysical possibility of zombies and inverts—or the view faces the quarantine problem.[[14]](#footnote-14)

# 6. Conclusion

The paradox of phenomenal judgment claims that epiphenomenal anti-physicalism entails phenomenal skepticism simpliciter on the grounds that it would prevent our experiential beliefs from depending on our experiences. As we saw, the paradox relies on content physicalism. And Chalmers challenges content physicalism by contrasting Mary and Inverted Mary. Yet the contrast between Mary and Inverted Mary supports *content-structure physicalism*:what we imagine is that Mary and Inverted Mary both think structurally equivalent thoughts that simply concern different sorts of color experiences (§4.1). And that leads to the quarantine problem. The contents of our pure phenomenal concepts depend on our experiences, but the rest of our psychology does not. Consequently, advocates of the paradox of phenomenal judgment can replace the paradox’s core premise with (1\*), the claim that our thoughts about how our experiences compare and contrast do not depend on whether those experiences actually compare and contrast.

There is, however, a more general lesson to be learned from the contrast between Mary and Inverted Mary. Our intuitions about consciousness are more complicated than we give them credit for. The received view in philosophy of mind is that thought experiments like Mary the color scientist, color spectrum inversion, and zombies offer strong prima facie support for epiphenomenal anti-physicalism. But this appearance is an artefact of our failing to consider the ways in which we can combine these classic thought experiments with one another. The contrast between Mary and Inverted Mary supports content-structure physicalism, and that is what ultimately leads to the epiphenomenal anti-physicalist’s problems with comparative phenomenal skepticism. And comparative phenomenal skepticism is a deeply counter-intuitive view. However, it would be strange to claim that Mary, inverts, and zombies provide prima facie support for physicalism. So, it seems that our intuitions about consciousness are not at ease with one another. And there simply may be no one metaphysical position that qualifies as the obvious champion of our intuitions about consciousness. Surprising as it may be, Mary and her (partially) color spectrum inverted and zombified counterparts not the friends of epiphenomenal anti-physicalism that they at first appear to be.

# Works Cited

Balog, K. (2012). In Defense of the Phenomenal Concept Strategy. *Philosophy and Phenomenological Research, 84*(1), 1-23.

Barz, W. (2017). Luminosity Guaranteed. *Pacific Philosophical Quarterly, 98*(S1), 480-496.

Block, N. (2006). Max Black's objection to mind-body identity. In D. Zimmerman (Ed.), *Oxford Studies in Metaphysics, II* (pp. 3-78). Oxford University Press.

BonJour, L. (2003). A Version of Internalist Foundationalism. In L. BonJour, & E. Sosa (Eds.), *Epistemic Justification: Internalism vs. Externalism, Foundations vs. Virtues.* Malden, MA: Blackwell.

Butler, J. (2011). Introspective knowledge of experience and its role in consciousness studies. *Journal of Consciousness Studies, 18*(2), 128-145.

Campbell, N. (2003). An Inconsistency in the Knowledge Argument. *Erkenntnis, 58*(2), 261-266.

Campbell, N. (2012). Reply to Nagasawa on the Inconsistency Objection to the Knowledge Argument. *Erkenntnis, 76*, 137–145.

Chalmers, D. (1996). *The Conscious Mind: In Search of a Fundamental Theory.* New York: Oxford University Press.

Chalmers, D. (2010). *The Character of Consciousness.* Oxford, UK: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780195311105.001.0001.

Dennett, D. (1991). *Consciousness Explained.* Boston: Little Brown and Company.

Dennett, D. (1995). The unimagined preposterousness of zombies. *Journal of Consciousness Studies, 2*(4), 322-326.

Dennett, D. (2005). *Sweet Dreams: Philosophical Obstacles to a Science of Cosnciousness.* Cambridge, MA: MIT Press.

Dennett, D. (2007). What RoboMary knows. In T. Alter, & S. Walter (Eds.), *Phenomenal Concepts and Phenomenal Knowledge: New Essays on Consciousness and Physicalism* (pp. 15-31). Oxford: Oxford University Press.

Duncan, M. (2017). Two Russellian Arguments for Acquaintance. *Australasian Journal of Philosophy, 95*(3), 461–474.

Ellis, J. (2010). Phenomenal character, phenomenal concepts, and externalism. *Philosophical Studies, 147*(2), 273-299.

Gertler, B. (2007). Introspecting Phenomenal States. *Philosophy and Phenomenological Research, 63*(2), 305-328.

Horgan, T., & Kriegel, U. (2007). Phenomenal epistemology: What is consciousness that we may know it so well? *Philosophical Issues, 17*(1), 123-144.

Jackson, F. (2003). Mind and Illusion. *Royal Institute of Philosophy Supplements, 53*, 251-271. https://doi.org/10.1017/S1358246100008365.

Jackson, F. (2004). Forward. In P. Ludlow, Y. Nagasawa, & D. Stoljar (Eds.), *There's Something About Mary.* Cambridge, MA: MIT Press.

Kirk, R. (1999). Why there couldn't be zombies. *Aristotelian Society Supplement, 73*, 1-16.

Levine, J. (2001). *Purple Haze.* Oxford and New York: Oxford University Press.

Nagasawa, Y. (2010). The Knowledge Argument and Epiphenomenalism. *Erkenntnis, 72*(1), 37-56.

Papineau, D. (2002). *Thinking About Consciousness.* Oxford: Oxford University Press.

Papineau, D. (2007). Phenomenal and Perceptual Concepts. In T. Alter, & S. Walter (Eds.), *Phenomenal Concepts and Phenomenal Knowledge: New Essays on Consciousness and Physicalism* (pp. 111-144). Oxford: Oxford University Press.

Shoemaker, S. (1975). Functionalism and Qualia. *Philosophical Studies, 27*(5), 291-315. https://doi.org/10.1007/bf01225748.

Shoemaker, S. (1996). *The First-Person Perspective, and Other Essays.* Cambridge: Cambridge Univsersity Press.

Silins, N. (2014). Experince Does Justify Belief. In R. Neta (Ed.), *Current Controversies in Epistemology* (pp. 55-69). London: Routledge.

Watkins, M. (1989). The Knowledge Argument Against the Knowledge Argument. *Analysis, 49*, 158-160.

1. Chalmers has an earlier response to the paradox of phenomenal judgment. See §5.2 for discussion. [↑](#footnote-ref-1)
2. Chalmers himself would likely agree with me on this claim, see fn. 7. [↑](#footnote-ref-2)
3. While Chalmers and I focus primarily on combinations of Mary the color scientist and color spectrum inversion, our arguments also work for combinations of Mary the color scientist and zombies. The main difference is that Zombie Mary would have experiential beliefs with gappy contents, which, depending on one’s theory of truth would make Zombie Mary’s beliefs either indeterminate or false. [↑](#footnote-ref-3)
4. Chalmers claims that standing phenomenal concepts tend to be more “coarse-grained” than the direct phenomenal concepts that they succeed (2010: 271). [↑](#footnote-ref-4)
5. See §5.2 for discussion of Chalmers’ earlier work on the paradox. [↑](#footnote-ref-5)
6. Blank canvas can represent periods of zombiehood. [↑](#footnote-ref-6)
7. While Chalmers does not explicitly discuss content-structure physicalism, much of what he says suggests the view:

Mary and Inverted Mary are physical twins, but they are phenomenally distinct, and this phenomenal distinctness…precisely mirrors their intentional distinctness (Mary believes that tomatoes cause R experiences, Inverted Mary believes that tomatoes cause G experiences).

…

There is a very strong intuition that the content of Mary’s [pure] phenomenal concept and phenomenal belief is determined by the phenomenal character of her visual experience in that it will vary directly as a function of that character in cases where that character varies while physical and other phenomenal properties are held fixed and also in that it will not varyindependently of that character in such cases. (2010, 266-7)

Here Chalmers suggests that the differences in Mary’s and Inverted Mary’s thought contents are circumscribed to their pure phenomenal concepts (and any complex concepts that include pure phenomenal concepts as component parts; 2010, 274-5) and concern the specific experiences that those pure phenomenal concepts were based on. This suggests that Chalmers endorses the claim that the overall structure or form of Mary’s and Inverted Mary’s beliefs will be identical so long as they remain physical/functional duplicates. [↑](#footnote-ref-7)
8. See fn. 13 for discussion of Chalmers’ similar ‘Dancing Qualia’ thought experiment from *The Conscious Mind*. [↑](#footnote-ref-8)
9. See §5.1 for discussion of a case of synchronic partial color spectrum inversion. [↑](#footnote-ref-9)
10. To avoid concerns that the thought components listed above are themselves linked to our experiences in something like the way that our pure phenomenal concepts are, we could instead use thought components like …*is the same property as…* and *…is a different sort of property from…*to formulate the quarantine problem. Chalmers claims that, “certain central concepts, such as causal and mathematical concepts, are not conceptually tied to phenomenal concepts” (2010, 275). And it seems plausible that our general concepts of property identity and property non-identity are not “conceptually tied to phenomenal concepts”. After all, they range equally well over phenomenal and non-phenomenal properties. For instance, if R and G are pure phenomenal concepts of phenomenal redness and greenness respectively, Mary could think, *R is a different sort of property from G. G is a different sort of property from being spherical*. *Therefore, R is a different sort of property from being spherical*. [↑](#footnote-ref-10)
11. Chalmers would likely agree that interactionist anti-physicalism and physicalism avoid the quarantine problem, since he argues that these views avoid problems like the paradox of phenomenal judgment for the reasons that I cite in the main text (2010: 293). [↑](#footnote-ref-11)
12. In §5.2, I discuss the possibility that phenomenal properties supervene on physical/functional properties with nomic necessity. [↑](#footnote-ref-12)
13. Chalmers’ dancing qualia case is a periodic color spectrum inversion scenario. He claims that the case is metaphysically possible and that this could raise “the worrying thought” that *we* are periodic inverts:

but that does not mean that it is plausible as an empirical possibility, any more than it is plausible that the world was created five minutes ago. As an empirical hypothesis, it seems far more plausible that when one’s experiences change significantly…one should be able to notice the change. (1996, 269)

Chalmers argument is difficult to interpret. Campbell, Jackson, Shoemaker, Watkins, and Dennett claim that epiphenomenal anti-physicalism entails that our experiential beliefs could not qualify as evidence about what our experiences are like. So, Chalmers cannot claim that our belief that we are not periodic color spectrum inverts serves as evidence that we are not periodic inverts without begging the question. Since the principle of organizational invariance is “an empirical hypothesis”, according to Chalmers, it is not provable apriori. So, Chalmers argument seems to be that there are non-evidential constraints on our theorizing about the laws of nature that license the inference that there is a principle of organizational invariance. Unfortunately, he does not say what those constraints are. Instead, he draws a comparison with the skeptical hypothesis that the world is five minutes old. But there is an important difference between the five-minute universe hypothesis and the hypothesis that we are periodic inverts that Chalmers does not address. The five-minute universe hypothesis is metaphysically possible on pretty much anyone’s accounting. So it is no one’s problem in particular. But only anti-physicalists grant the metaphysical possibility of periodic color spectrum inversion. So, Chalmers cannot be claiming that the metaphysical possibility that we are periodic color spectrum inverts is everyone’s problem and so no one’s problem in particular. But, since Chalmers’ does not say what the relevant principles of inference are, it is difficult to offer a definitive assessment of his argument. This is why I offer responses in the main text that do not require me to dispute the existence of the principle of organizational invariance. [↑](#footnote-ref-13)
14. Similar comments apply to *Russellian monism*, the view that phenomenal properties (or proto-phenomenal properties) are the intrinsic properties that underlie the structural and dynamic features of the particles, waves, and fields described by fundamental physics. [↑](#footnote-ref-14)