

## Introduction

*Note 1: This post is a collaboration between Sebastian Montesinos, Joseph Lawal, Lucas Collier, and Benjamin (Truth Teller). Sebastian & Lucas are co-authors of Naturalism Next and their info can be found on the [about us page](#). Joseph is a graduate student in philosophy at Yale University studying the intersection of philosophy of language, philosophy of mind, and epistemology. Benjamin (who goes by the moniker Truth Teller) is a philosophy autodidact and friend of the blog, whose work can be found [here](#).*

*Note 2: This post is long, and blogger is bad with hyperlinks. Therefore, we are providing a version of this piece you can access on google docs with a table of contents that will let you jump to specific sections, for those who would prefer this option.*

Recently, Brian Cutter & Dustin Crummett published a preprint of a paper presenting a new argument for theism based on ‘psychophysical harmony’ that has exploded in the online apologetics sphere. Amongst the claims that have been made about this paper is that it is the ‘strongest argument for God in history’, that ‘atheists can’t beat it’, and that, in general, the probability judgment present is so strong as to utterly swamp any atheistic evidence to the contrary. In this piece, we present a critical analysis of the argument, and conclude (unsurprisingly) that these assessments are false. In fact, ‘psychophysical harmony’ is no serious threat to naturalists at all. The objections we present in this piece are mostly novel, undiscussed problems that undercut the core mechanics of the argument and are therefore, we think, very important contributions to the debate.

Our response is divided into two broad sections. *First*, we present concerns we have about the dialectical force of the argument. We point out that the advocates of psychophysical harmony do not take as seriously as they should the considerable number of views that avoid their argument, even granting all of its assumptions. *Second*, and more importantly, we present six positive arguments against the core claims and assumptions of the psychophysical harmony argument that, if correct, completely undercut its effectiveness. Notably, the assumptions that we highlight in this second section demonstrate that there are many views in the mainstream of philosophy that undercut psychophysical harmony. Therefore, it is not only those brave souls who want to embrace some of the radical positions outlined in the first sections of our piece who can avoid the argument: most naturalists can do so happily.

Importantly, two parallel aims run throughout this piece. First, we aim to challenge on philosophical grounds the claims being made by Cutter & Crummett in their arguments. Second, we aim to make explicit to advocates of the argument the assumptions of which Cutter &

Crummett are doubtless already aware. This second aim is important insofar as the argument from psychophysical harmony has, in our view, had an outsized impact on both theists and atheists who are not aware of the various ways in which the argument presupposes the truth or falsity of highly contested positions.

To keep this post as short as possible, we have not provided an overall summary of the argument from psychophysical harmony. [The paper to which this post is responding is available online.](#)

## Table of Contents

[Introduction](#)

[Table of Contents](#)

[Concerns](#)

[Views that Avoid Psychophysical Harmony](#)

[Type A Physicalism](#)

[Liberal Naturalism and Kripkean Conceivability \(Davidson, Putnam and Kripke\)](#)

[Eliminative Materialism](#)

[Error Theory and Moral Anti-Realism](#)

[Axiarchism and Directed Naturalism](#)

[Why These Alternative Views are Serious](#)

[Objections](#)

[Objection #1: A Dilemma for Normative Harmony](#)

[Introduction](#)

[Horn #1: Taking phenomenal introspection \(very\) seriously](#)

[Horn #2: Dispelling phenomenal introspection, or, pain's normative status as contingent](#)

[A Parallel to a Real Case](#)

[Conclusion](#)

[Objection #2: Semantic Harmony and Content Individuation](#)

[Objection #3: Disharmonious Doubles: Liars, Lunatics, or Linguistically Confused?](#)

[Introduction](#)

[Horn #1: True](#)

[Horn #2: Lying](#)

[Horn #3: Mistaken](#)

[Final Considerations](#)

[Objection #4: Causation and Explanation](#)

[Objection #5: Understated Evidence](#)

[Introduction](#)

[Our World is Much More Disharmonious than You Think It Is](#)

[Why Cutter & Crummett's Response to Understated Evidence is Insufficient](#)  
[Objection #6: The Revenge Problem](#)  
[Introduction](#)  
[Counter #1: Metaphysical Necessity](#)  
[Counter #2: Simple God](#)  
[Counter #3: Simple Laws](#)  
[Counter #4: Classical Theism](#)  
[Works Cited](#)  
[Conclusion](#)

## **Concerns**

As mentioned in the introduction, we think the advocates of psychophysical harmony do not take the views that avoid their argument as seriously as they should: they fail to mention some, and of those they do mention, they downplay the importance of these views for the dialectical significance of their argument. This is sometimes true of Cutter & Crummett, and certainly true of the apologists weaponizing the argument. We begin by mentioning some of the views that rob the psychophysical harmony argument of much or all of its force, and then explain why the existence of these views means that psychophysical harmony cannot be said to be a seriously effective argument for theism.

### Views that Avoid Psychophysical Harmony

*By: Sebastian Montesinos, Lucas Collier, and Joseph Lawal*

#### *Type A Physicalism*

Famously, Kripke's *Naming and Necessity* ends with a conceivability argument against identity theory. He writes the following about God creating humans:

It would seem, though, that to make the C-fiber stimulation correspond to pain, or be felt as pain, God must do something in addition to the mere creation of the C-fiber stimulation; He must let the creatures feel the C-fiber stimulation as pain, and not as a tickle, or as warmth, or as nothing, as apparently would also have been within His powers. (p. 154.)

Rather than just have us imagine zombies who experienced no phenomenal state whatsoever because of C-fiber stimulation, Kripke asks us further to imagine the possibility of

humans who, upon the firing of the C-fibers, experience different phenomenal states entirely. Although they come to it as filtered through David Chalmers' work, this scenario is the heart of Cutter & Crummett's argument from psychophysical harmony. They ask us to conceive of several scenarios wherein behaviors, and functional and physical facts remain the same as they are in the actual world, but the correlating phenomenal states are altered.

The possibility and conceivability of these situations depends on what the nature of the correlation between physical-behavioral states and phenomenal states is. Cutter & Crummett explicitly rely on a nomological correlation between the two: it is the addition of a "psychophysical law" to a physical-behavioral state that allows the deduction of the correlated phenomenal state. To them we are lucky to be in a world with psychophysical laws such that our physical-behavioral states seem to match so well with their phenomenal correlates. For example, C-fibers firing and aversive behavior match well with the distress and suffering of pain, but not as well with the joy of a good joke. Cutter & Crummett argue that theism (or similar hypotheses—more on that later in this section) predict that the psychophysical laws would be such that physical-behavioral facts seem to match so well with their correlated phenomenal states.

We illustrate in the remainder of this section that there are serious views that reject our need for a psychophysical law in the deduction of a phenomenal state from physical-behavioral one, and instead posit that the physical-behavioral state by itself sufficient for the deduction of the correlating phenomenal state. In terms of Kripke's thought experiment, these views are tantamount to stating that when creating the universe, God needn't fix particular phenomenal states to physical-behavioral ones. Rather, the work was done for Him when he created those physical-behavioral states.

The first example is Dennett's analytic functionalism. His view is complex and elaborated extensively throughout the course of his career, but I will focus on a few relevant points. I will let Dennett state the first:

...consciousness is not a single wonderful **separable** thing ('experiential sensitivity') but a huge complex of many different informational capacities that individually arise for a wide variety of reasons.... (Dennett 1996, p. 324, emphasis added.)

This quote demonstrates a difference in attitude between Cutter & Crummett, and Dennett: Cutter & Crummett treat phenomenal experience as a conceptually distinct entity that can be divorced from its correlated physical-behavioral states and then 'dragged and dropped'

or swapped out for others. Dennett instead offers a view of consciousness that is not so neat which contends that consciousness is just a range of capacities and behaviors and therefore is not cleanly separable from physical-behavioral states and functions. Because of this, Dennett calls zombies not just inconceivable, but 'preposterous' because, crudely, once we suppose that these zombies mirror us identically in all of their physical characteristics, capacities, and behavioral dispositions, there is nothing left to distinguish them from us.

This would carry over similarly to instances of disharmony: if in a disharmonious world our physical characteristics, capacities, and behavioral dispositions are identical, there is no room left for a difference in or inversion of phenomenal states.

This is true not just of Dennett's analytic functionalism, though. Any view which rejects a conceptual or epistemic gap between physical-behavioral facts and phenomenal states has available to them almost the same line of reasoning. The form of this objection can be refined as follows: Cutter & Crummett, when laying out an example of disharmony, start by noting that in the actual world some physical-behavioral state *S* correlates with the phenomenal state *M*. Then they ask us to imagine another world *w* where *S* obtains but *M* is swapped with a drastically different counterpart. Views that reject the epistemic gap (e.g. *a priori* physicalism, illusionism) hold that there is a necessary entailment relation between *S* and *M*. Thus it is a conceptual truth that in *w*, it cannot be the case that both *S* and  $\neg M$ . This marks for these views the inconceivability of *w* and the disharmony instances. Cutter & Crummett explicitly acknowledge that their argument presupposes the falsity of *a priori* (or "type A") physicalism - that is, they are clear about presupposing an epistemic gap between phenomenal and physical/functional truths (p. 30). Dennett is one of several *a priori* physicalists, and presents one way of denying that the argument from psychophysical harmony can get off the ground at all.

### *Liberal Naturalism and Kripkean Conceivability (Davidson, Putnam, and Kripke)*

Many philosophers find *a priori* physicalism, especially given its tendency to lead to reductionism and eliminativism, too radical and therefore unacceptable. This might explain in part why Cutter & Crummett flag their presupposition but do not defend it. This is fine as far as it goes; it is often useful in philosophy to explore conceptual space by taking for granted even a highly controversial position (and, as Cutter & Crummett note, *a priori* physicalism is a minority position). Of course, this assumption still weakens the dialectical appeal of their argument, since their argument has nothing to say to the *a priori* physicalist, and leaves open the possibility that

the opponent of their argument need only adopt a priori physicalism to avoid accepting an unwelcome conclusion. But Cutter & Crummett are likely to see *a priori* physicalism as too radical to pose a serious threat to their argument. Again, *a priori* physicalists tend toward reductionism and eliminativism, and many people struggle to take reductive or eliminative accounts of the mind seriously.

It bears remembering, then, that one need not be a reductionist or eliminativist to question the conceivability of zombies (that is, to question the epistemic gap upon which the argument from psychophysical harmony depends). To briefly mention some alternatives: liberal naturalists firmly oppose the reduction of mental talk to physical talk. They affirm the importance of scientific explanations, but also deny that all explanations are ultimately reducible to scientific ones. But liberal naturalists like Donald Davidson and Hilary Putnam (at least, the Putnam of *The Threefold Cord*) will also deny the epistemic gap. Donald Davidson's anomalous monism identifies token mental states with token physical states, but rejects out of hand the possibility of science of psychology modeled after physics. Davidson takes seriously our ordinary talk of beliefs, desires, and other mental states without allowing that there is an epistemic gap between such states and the physical states to which they are identical. Similarly, Putnam argues in *The Threefold Cord* that the very notion of a *correlation* between mental and physical states is confused. However, he also argues vigorously against reductionism and epiphenomenalism. (It may also be worth noting, though we have not explored it here due to a desire to keep this post a manageable length, that Putnam's Quinean rejection of absolute *a priority* would also likely neuter the argument from psychophysical harmony. Quinean naturalists have strong reason to reject - or, like Putnam, seriously weaken - the notion of *a priority*, which would give them reason to deny that the kind of *a priori* analysis on which the argument depends is coherent.)

One can even question the epistemic gap without committing to a theory of mind at all. On Kripke's account of conceivability (in "Identity and Necessity" and *Naming and Necessity*), we are often wrong about what is conceivable. It might have seemed to us, before the advent of modern chemistry, that water's turning out to be H<sub>2</sub>O and its turning out to be XYZ are equally conceivable. However, to Kripke, given that water actually is H<sub>2</sub>O, it is not only impossible that water be XYZ, it is also inconceivable. When we take ourselves to be conceiving of water as XYZ, we aren't conceiving of water at all, but an epistemic counterpart to water. Kripke puts this notion of conceivability to use in arguing against identity theory, but one could use it equally powerfully to raise doubts about whether states of affairs – especially complex states of affairs involving fraught scientific and philosophical notions like brain states and phenomenal states –

are genuinely conceivable, or how far we should trust their apparent conceivability. On Kripke's account, it's *easy* to be wrong about what is conceivable, since what is conceivable is tied so closely to what is possible, and it is not easy to tell what is possible. Yablo (in "Textbook Kripkeanism and the Open Texture of Concepts") uses Kripke's work to call into question the conceivability of zombies in something like the way I have just suggested.

The fact that one can call into question the existence of an epistemic gap between physical and phenomenal truths even *without* committing to reductionism or eliminativism means that the dialectical weakness in the argument from psychophysical harmony is perhaps more significant than Cutter and Crummett have admitted, and certainly more significant than many newfound adherents of the argument have realized.

### *Eliminative Materialism*

Another position in the philosophy of mind that Cutter & Crummett do not mention is eliminative materialism. According to eliminative materialism, our commonsense understanding of the mind that references propositional attitudes is radically mistaken, and will eventually be displaced by a future, superior conceptual framework inspired by neurosciences. If eliminative materialism is true, then there are no beliefs, and so it no longer makes sense to talk about whether there is harmony between our beliefs and our phenomenal states. According to eliminative materialists, 'justification' as it has been classically conceived in terms of the relationship between judgements and perceptual states will require serious reform. Of course, one could suggest that in the new framework eliminative materialists propose there will be some kind of harmony between the reformed mental states and epistemology proposed by eliminative materialists and our phenomenal states. However, until such time as that framework has been proposed it would not be reasonable to prejudge that it will turn out in a way that favors an argument from psychophysical harmony. Of course, eliminative materialism is a very radical position, and not many hold to it, so it is not surprising that Cutter & Crummett do not mention it. Nevertheless, the fact that it is radical is not a reason to reject it out-of-hand, and it is yet another option naturalists have when confronted with psychophysical harmony [1]. Notably, at least two of the authors of this post are sympathetic to it.

## *Error Theory and Moral Anti-Realism*

Of the views that could cause problems for Cutter & Crummett's argument, this is the view they dedicate the most sustained discussion to. If normative error theory is true, then there are no normative facts, so there is no normative role for pain to play that could be harmonious with its functional role. A parallel point is true for epistemic error theory. There may be ways to state the psychophysical harmony argument such that this issue is skirted, at least in the case of normative error theory [2]. Nevertheless, since Cutter & Crummett themselves think this position does undermine normative harmony, and because what they say about it is instructive for how we ought to respond to more radical theories that might cause problems for psychophysical harmony, it is worth examining what they say.

Cutter & Crummett respond to error theory as follows:

"It seems about as self-evident as anything in philosophy that excruciating pain is bad, or that experiences justify certain beliefs. If denying these claims is the price of upholding naturalistic atheism, the price is too high" (p. 10).

No error theorist will feel very threatened by this. First, the way Cutter & Crummett have phrased this response trades on the ambiguity between 'stance-independent badness' and 'badness'. The error theorist might be very happy to agree that pain is bad in the sense that they disapprove of pain, or dislike pain, or that pain otherwise conflicts with their preferences and values. What the error theorist will deny is that 'pain is bad' is a stance-independent moral fact. Second, while moral anti-realists are certainly in the minority of philosophers, metaethical anti-realism enjoys a healthy following amongst naturalists, and has many motivations quite independently of any motive to undermine psychophysical harmony. It is important to note that under the umbrella of anti-realist positions that cause problems for normative harmony, moral error theory, which is a cognitivist version of anti-realism, is just one possibility. Any non-cognitivist, in virtue of denying that moral statements are truth apt, would surely deny that there is any true proposition about the proper normative role that states like pain should play [3]. While there are more moral realists than anti-realists in philosophy, it can hardly be alleged that anti-realism is so ridiculous as to be self-evidently false: roughly a fourth of philosophers identify as such [4]. Cutter & Crummett are more justified in asserting that epistemic error theory is quite a radical and fringe theory, but it is still a view worth mentioning, and we take it that no epistemic error theorist will either be threatened by the implication that their view is terribly radical either (see again footnote 1).



### *Axiarchism and Directed Naturalism*

Another set of views that avoid the psychophysical harmony argument for very different reasons than the views discussed so far is a kind of directed naturalism or axiarchism. Cutter & Crummett explicitly discuss these kinds of views. On these views, the universe is directed towards producing value-conferring states in some way. Thomas Nagel was famously a proponent of this kind of view, and the panpsychist Philip Goff also takes an approach in this spirit to the problem posed by psychophysical harmony. Since proponents of these views hold that the universe is directed towards some combination of meaning, purpose, value, and goodness, the probability that we would observe psychophysical harmony on these views is high. Therefore, unlike the views discussed so far, these theories do not deny harmony or its assumptions in some way. Instead, this approach fully accepts the existence of harmony but offers a view that accounts for it as well as theism.

This is, it should be noted, a very large umbrella: all kinds of “theism-adjacent” (p. 2) views explain psychophysical harmony but do not entail the existence of anything like the God of classical theism or orthodox Christianity. Cutter & Crummett are under no illusions in this regard, but some advocates of the argument from psychophysical harmony seem to have overlooked the fact that the argument gets you God in some robust sense only as one possibility amongst a large number of alternatives. It is, in fact, perhaps misleading to label these alternatives as “theism-adjacent,” since many of them (such as Nagel’s view and perhaps Goff’s) are best labeled as overtly atheistic, even though they posit some kind of value-orientation in the universe. The presence of these alternatives means that the argument from psychophysical harmony can never constitute, on its own, anything like an argument for the existence of a perfect being; it is an argument for a wide range of views. The existence of a perfect being is one way amongst many to explain psychophysical harmony, and so any Christian apologist employing the argument would need to rely on an additional argument (some of which apologists apparently take the argument from psychophysical harmony to replace) to take the additional step to perfect-being theism, classical theism, or Christian orthodoxy.

In fact, it is likely that these views account for the data *better* than theism for reasons we will elaborate on much later in the post on the section on understated evidence—what we actually observe is a universe weighted towards harmony but with significant disharmony, which fits better with a ‘generally weighted towards the good’ theory than a ‘perfect being’ theory.

---

[1] In these sorts of dialectical situations, it is valuable to heed what Michael Devitt says about eliminativist positions: “It is futile to keep emphasizing how shocking eliminativism is and how difficult it would be to live with. Your average eliminativist is thoroughly aware of this; indeed, she enjoys being an *enfant terrible*. And she is fond of pointing out how often in the history of science the truth has been shocking. This is a very good point. What eliminativism calls for in response is neither transcendentalism nor cries of horror, but arguments aimed at what the eliminativist actually says.” (Devitt, 1990)

[2] To see how, consider the parallel move made in discussions around the problem of evil: in response to atheists, the theist asserts that if naturalism is true then ‘good’ and ‘evil’ have no meaning or there are no true moral statements, and so the atheist cannot make claims about the evidential weight of ‘evil’. One good response the atheist should make to this claim is to point out that the data in question can be stated neutrally as to avoid any metaethical commitment just by referencing facts about pain, suffering, the lack of wish fulfillment, etc. If theism is true, because God is all-good, these facts are unexpected, but not if naturalism is true. Similarly, in the case of the psychophysical harmony argument, the advocate might be able to state the data not as being that pain and pleasure serve their proper normative roles—a fact that antirealists will deny—but instead that pain and pleasure are correlated with certain behaviors. Then, they can say that on theism, because God is all-good, we should expect these correlations, while on naturalism, we have no reason to.

[3] Interestingly, non-cognitivism seems more resistant to the alternative framing to the psychophysical harmony argument suggested in footnote 2, because a committed non-cognitivist may have problems entertaining a conditional on which they need to imagine that cognitivism is true. Nonetheless, an adjustment to the argument that frames the data as between particular phenomenal states, functional states, and attitudes that God has might still work.

[4] Bourget and Chalmers 2020

### Why These Alternative Views are Serious

*By: Sebastian Montesinos*

So, there are many positions that naturalists might adopt that deprive the psychophysical harmony argument of some or all of its force. It is worth noting that none of these positions could rightly be called *ad hoc* since all of them were formulated far before the psychophysical harmony argument was conceived, and have ample independent justification.

We now examine what the dialectical implications of these positions are. We conclude that in the strong sense of 'effective' in which an argument can move someone's credence such that they adopt a new position, that psychophysical harmony is inert: almost no naturalists should become theists due to psychophysical harmony [1].

We are extremely confident that most naturalists and atheists, prior to considering the psychophysical harmony, would assign a much higher prior probability to most if not all of the positions just listed than theism. Most theists seriously underestimate how much of an anathema theism is to the way that naturalists understand the world. Shifting from naturalism or atheism to theism is a much more radical shift in terms of the commitments and changes it requires to one's web of beliefs than shifting from naturalism to type A physicalism, liberal naturalism, axiarchism, etc. Therefore, if a naturalist is presented with psychophysical harmony and doubts no part of the argument, it will almost certainly be more rational for them to remain a staunch atheist, and simply adopt one of the many naturalistic or non-naturalistic hypotheses consistent with atheism on which psychophysical harmony is not a problem [2]. Of course, Cutter & Crummett could argue that every one of these views is false on the basis of other arguments, but that would make psychophysical harmony *as an argument for theism* reliant on the success of highly controversial arguments in philosophy of mind, metaphysics, metaethics, philosophy of language, and the philosophy of religion, which is something Cutter & Crummett seem to want to avoid. In conclusion, insofar as arguments are meant to convince one's interlocutor of a position, psychophysical harmony appears to have little to no dialectical force on its own, *as an argument for theism*.

However, while Cutter & Crummett ought not downplay the importance of these alternative ways to avoid psychophysical harmony, they nonetheless do not render the argument totally ineffective given a weaker sense of 'effective'. First of all, if psychophysical harmony rules out certain views for naturalists, that is an extremely valuable insight for naturalists attempting to construct an explanatorily sufficient worldview. Secondly, if naturalists do need to shift to *specific, controversial theses* in order to motivate their rejection of psychophysical harmony, that does not necessarily mean that naturalists have no evidential problem with respect to the data. This is because the way to evaluate the effect of an auxiliary

hypothesis that is uniquely added to naturalism to predict the data, in a Bayesian sense, is using the weighted average principle as follows:

$$\Pr(\text{PH}|\text{N}) = \Pr(\text{PH}|\text{N}\&\text{A}) + \Pr(\text{A}|\text{N}) + \Pr(\text{PH}|\text{N}\&\sim\text{A}) + \Pr(\sim\text{A}|\text{N})$$

What the weighted average principle says, in plain english, is that expansions to naturalism need to both predict the data in conjunction with naturalism *and be likely on naturalism antecedently*. And, given that the expansions naturalists will need to adopt are sometimes quite controversial, arguing that  $\Pr(\text{A}|\text{N})$  is very high will not be easy. Of course, given that there are a number of possible hypotheses naturalists can adopt to explain the data, all naturalists need to show is that  $\text{A}_1 \vee \text{A}_2 \vee \text{A}_3 \dots$  is high, which is an easier task, but not *obviously* viable. Thus, even if psychophysical harmony does not work as an argument for adopting theism, it may still show that we ought to assign a much lower credence in naturalism than previously thought, and thus assign a correspondingly higher credence in theism (note, though, that the naturalist could just as easily and probably more happily shift their credence to some non-naturalistic but atheistic view that explains harmony). Thus, we conclude that the naturalist who wants to strike down the psychophysical harmony without any cost to naturalism or without an argumentative burden has more work to do. We now turn to doing that work.

---

[1] Note that here the term ‘theism’ can be construed either narrowly to mean omnitheism or more broadly to mean any kind of fundamentally, value-directed view and what we say in the following sections will still apply. We mention this because while the less careful apologists have simply been framing psychophysical harmony as an argument for omnitheism, Cutter & Crummett are careful enough to sometimes frame it as an argument for more generally value-directed hypotheses.

[2] We suspect this will be true for many agnostics as well. Indeed, while we did not examine this issue in depth, there are many non-naturalistic views on which psychophysical harmony is explained that do not succumb to the evidential challenges that orthodox theism does such as the problem of evil.

## Objections

## Objection #1: A Dilemma for Normative Harmony

*By: Sebastian Montesinos*

### *Introduction*

Our argument against normative harmony has two horns. On the first horn, we take phenomenal introspection seriously, granting that we can “see” the normative content of states such as pain and pleasure from introspection. On this horn, in virtue of taking introspection seriously, we also have strong reason to think that we cannot epistemically divorce pain from its functional profile, which undercuts Cutter & Crummett’s argument. On the second horn, we do not take phenomenal introspection seriously: on this horn, while we can easily divorce pain from its functional profile, it looks like pain’s normative status is now contingent on its functional profile, and that the natural counter to this claim is undercut. To the extent that horn #1 is accused of illicitly relying on our knowledge of the psycho-functional correlates in this world, the same applies to the theist’s claim about pain’s normative status in horn #2.

### *Horn #1: Taking phenomenal introspection (very) seriously*

One response to normative harmony would be to suggest that pain is necessarily epistemically linked to reports of displeasure and agony, avoidance behavior, and other parts of its behavioral profile. In other words, it is not possible to imagine pain without this accompanying functional profile of pain. We have already seen a number of views on which we ought to take this possibility seriously, and it is crucial to note that one does not need to commit to a hard illusionism about consciousness on which pain just *is* its functional profile to adopt this view. It may very well be that there is some intrinsic phenomenal content to pain separate from its profile, but that the two are *necessarily linked*, such that imagining one requires imagining the other. Nor does it commit one to thinking that we can deduce what it is like to experience pain *a priori*. Even if we know that pain must be conceptually linked to a particular profile of behaviors and responses, knowing what it is like to experience pain is not a kind of propositional knowledge you acquire from learning the physics, it is a kind of prelinguistic knowledge by acquaintance you can only acquire by having your synapses configured in the right way, and having the right neurological signal propagate through your brain. Therefore, this response does

not rely on a strong illusionism about consciousness that many people find implausible (though, it does offer an interesting sort of middle position between type A and type B physicalism).

Why should anyone accept this view? As already mentioned, there are quite a number of physicalist accounts of consciousness that entail this view, including non-illusionist and non-reductive accounts. Additionally, there are purely conceptual reasons to accept this view: our very concept of pain rests on its relation to certain behaviors. Out of curiosity, I surveyed a few philosophy undergraduates about their intuitions in these cases, and all of them thought that in separating pain out from these behaviors you would no longer be talking about pain. (Perhaps, even though it is the opposite of the view he holds, this is related to Kripke's intuitions regarding conceivability.) We will also later see how thinking that there is some way to conceive of pain having inverted functional roles relies on a kind of epiphenomenalism+ that falls afoul of serious problems regarding causation and explanation (Objection #4).

However, let us focus on trying to levy the very thing Cutter & Crummett use to motivate their understanding of pain as bad in the paper, phenomenal introspection, in favor of this horn. Personally, I do not think phenomenal introspection is a coherent idea (more on this soon). However, let us suppose, for a moment, that it is. It seems as true as anything that we could know introspectively that when you introspect on the feeling of what it is like to be in pain, you can just see that this feeling must be accompanied by a particular functional profile. One might claim that this illicitly relies upon our knowledge of the psycho-functional correlates in this world. In other words, in learning this fact from introspection, we've actually relied on our background knowledge about psycho-physical correlates. However, if we are taking phenomenal introspection seriously, this is the kind of process wherein we can *just see*, without reference to our intentional states or functional correlates, certain truths about pain. More importantly, *if* this objection is correct, it undermines the most plausible counter one might give to the second horn of our response, as we shall soon see:

#### *Horn #2: Dispelling phenomenal introspection, or, pain's normative status as contingent*

The second response to normative harmony is the following: pain's normative status is contingent upon the world we are in. In a world where pain and pleasure are inverted, in virtue of pain's functional profile being such that we seek it out and report that it is good and enjoyable, pain is no longer intrinsically bad. One reason to resist this is to point out that our phenomenal introspection seems to strongly reveal that pain is stance-independently bad. By introspecting, we can just see that pain is the sort of thing that is intrinsically awful. However, this horn of the

argument does not take phenomenal introspection so seriously. The very idea that we can phenomenally introspect pain's badness is confused: the intuition or judgment that pain is bad is a *second-order* judgment *about* our phenomenal states.

Indeed, I would go farther and suggest that I do not even understand what it means for pain's stance-independent badness to be solely a part of one's phenomenal introspection. Phenomenal introspection is a pre-linguistic, non-propositional kind type of introspection that cannot import the kind of propositional meaning we need to learn that something is stance-independently bad. That can only be achieved by going through the representation process of forming a propositional attitude, which must necessarily reference the folk psychological laws and concepts that attitude is embedded in. This web of laws and concepts is *contingent* on the world that we live in: for instance, consider the familiar homily that if someone believes that they are in pain and desires not to be in pain that they are going to avoid pain. This relies on our understanding of the functional profile of pain in this world. These homilies would be different in the worlds in which pain's functional profile is inverted. The psychophysical harmony argument asks us to prejudice the phenomenal judgments we make in this world against those worlds, extrapolating our intuitions in this world as applying in other ones, such that pain is intrinsically bad across worlds. *From the perspective of those other worlds* in which pain has a functional profile such that we seek it out and judge that it is good, the intuition here evaporates and inverts. Additionally, it also illicitly relies on our knowledge of the psycho-functional correlates in this world, since, if I am right, that pain is stance-independently bad cannot be *derived solely* from phenomenal introspection. Instead, that pain is bad must be a judgment we make on the basis of the psycho-functional correlates of pain.

Cutter & Crummett do have an argument against this general approach:  
"Consider a disembodied Cartesian mind whose overall experience is exactly the same as that of the torture victim. The experience of the disembodied mind has no causal ties to any physical behavior or physical stimuli (because she has no body or physical sense organs). Still, it seems self-evident that her experience is bad, one she has reason to avoid. We submit that it is not even conceivable that there should be a perfect phenomenal duplicate of the torture victim whose experience is not intrinsically bad." (Pg. 11-12)

Note that this passage relies on us *taking phenomenal introspection seriously*, such that we can see, solely from our phenomenology, that pain is self-evidently, stance-independently bad across worlds. I have argued that this is not possible, and that our bias towards viewing such a thing as self-evidently bad *relies illicitly on our knowledge of the functional profile of pain in this world*. It would be manifestly unreasonable to think that judgments that are contingent on

a conceptual framework that relies on the functional correlates of states like pain *in this world* would still be true when you change those correlates, and so thereby change the structure of that conceptual framework. In the world where pain's functional profile was inverted, it would be intuitively just as inconceivable that the torture victim's pain was bad. And, insofar as we do think phenomenal introspection *is* something that we can divorce from this greater profile of behavior and second order judgements, *it is just as obvious that phenomenal introspection reveals that pain must be conceptually linked to a particular functional profile as it is obvious that pain is stance-independently bad.*

### *A Parallel to a Real Case*

One interesting way to think about the above dilemma is to look at its parallel to something that actually exists in our world: pain asymbolia. In this condition, patients appear to have the same phenomenal experience of pain as other organisms, but without any outward indication of agony, reports of displeasure, and avoidance behavior. One natural kind of response to this condition is to say 'that isn't really pain', in that pain as a concept must be linked to these various functional outputs. Once you divorce it from those outputs you are no longer really talking about pain, but something else. This would be analogous to accepting horn #1 in the above dilemma. The other option would be to agree that pain asymbolia is still pain, given that people report having the same phenomenal experiences, but to argue that pain must be divorced from the 'intrinsic badness/awfulness' that we usually associate with it. In other words, the 'badness' of pain comes not from the raw phenomenal experience of pain, but from its associated functional profile, making its badness contingent on that profile. This would be analogous to accepting horn #2 of the above dilemma. What does not look reasonable is to maintain both that pain asymbolia patients are really experiencing pain, and that pain is always stance-independently bad regardless of its functional profile, which is what the psychophysical harmony advocate's positions entail

### *Conclusion*

To sum up, normative harmony faces an intractable-looking dilemma. If pain being bad is a second-order judgment we make using an inherited conceptual framework, then it is a judgment that *necessarily* references the contingent functional profile of pain in this world. If this is true, it would be unreasonable to think that these same judgments would hold in the worlds in which the 'laws' underlying our conceptual framework were completely different or inverted.



Alternatively, perhaps pain being bad is something we have some sort of direct access to in the absence of any conceptual framework through phenomenal introspection. Phenomenal introspection is, I think, an inherently confused concept *and* there are very powerful reasons to think that our judgements and beliefs must be theory-laden posits that reference an implicit conceptual framework that we all inherit. Nonetheless, if we do have direct access to our phenomenal states, it is just as obvious that pain is bad as it is that for pain to be pain it must be followed by a certain functional profile of attitudes and behaviors, but the psychophysical harmony advocate cannot accept this.

I do not deny that there are ways one could try to resist this dilemma. However, not only is it the burden of the psychophysical harmony advocate to formulate this, it is clear that this objection establishes a way out of psychophysical harmony that is far more serious than simply adopting a radical position: the views relied on here are neither unreasonable nor fringe.

#### Objection #2: Semantic Harmony and Content Individuation

*By: Joseph Lawal*

The argument from semantic harmony begins with the following claim:

In many cases, the psychophysical laws pair phenomenal states with physical states in a way that generates a semantic correspondence between our judgments/reports and our phenomenal states. (p. 13)

The argument amounts to this: We make judgments (or at least we *can* make judgments) like “I am having a reddish experience.” Often these judgments or the reports that we make as a result of the judgments are true. This is lucky because we might instead have made these *same* judgments or reports but without the phenomenal states which correspond to and make true our actual judgments and reports. That is, in an enormous number of conceivable worlds, these judgments and reports might have been false, and so it is lucky that the judgments and reports we make are true.

This is, it seems to me, much too quick. Cutter & Crummett assert that psychophysical laws “generate” a semantic correspondence between judgments and reports on the one hand and phenomenal states on the other. But why should we think about the individuation of the contents of our judgments and reports this way? If, for instance, we take seriously content externalism, then it could very well be that our judgments and reports about phenomenal states

*constitutively depend* on phenomenal states they are about. There's no "generating" work for psychophysical laws to do. Semantic harmony is presented as a problem for non-theists without more than a cursory nod to accounts of the individuation of semantic and representational mental content which raise very serious questions about the coherence of the scenarios Cutter & Crummett describe.[1]

The suggestion here, to reiterate, is that judgments and reports about phenomenal states may constitutively depend on the phenomenal states they are about.[2] We can refer to the familiar Twin Earth cases[3] to understand how and why representational content is individuated anti-individualistically. It is widely accepted [4] that on earth, Oscar, when he thinks, "I need to drink some water," is having a thought about *water*, and when Twin Oscar on Twin Earth, where the lakes and rivers etc. are full of XYZ instead of H<sub>2</sub>O, thinks "I need to drink some water," he is instead having a thought about *twin water*, which occupants of Twin Earth call "water" but which, because it has a different microphysical constitution from what we call "water" (H<sub>2</sub>O), we would not call "water.". Oscar's and Twin Oscar's brain states are identical but they are having different thoughts because their social and physical environments differ. This point generalizes to judgments and reports, i.e., when Oscar judges "this is water," he is making a different judgment than Twin Oscar when he judges "this is water."

It is not hard to see how this bears on the question of psychophysical harmony. Cutter & Crummett suppose that in a world in which I never have any reddish experiences, I can still make the judgment "I am having a reddish experience," but that judgment will be false. But does reflection on Twin Earth cases not suggest that in the scenarios Cutter & Crummett describe, I cannot form the relevant judgment at all? The argument from psychophysical harmony appears to presuppose that, say, if I imagine a world in which water does not exist, I can imagine myself making judgments about water in that world. But a strong case could be made that this is not so.[5] Perhaps we should say instead that I would not be making the same judgment (i.e., I am having a reddish experience), a judgment which happens to be false at that world, because that judgment *constitutively* depends on the phenomenal state it is about; take away the state, and you destroy *that* judgment. Judgments of similarity will also disappear – it is not the case that my judgment "this experience is more similar to that experience than this other experience" (p. 16) is *false* in most of these worlds, but that I fail to make the relevant judgment at all.[6] We are not, after all, thinking about what it might mean for us to describe the relevant scenario using

our language; we are after what the words would mean *in the mouth of someone in the relevant counterfactual*.

Notice that even if Cutter & Crummett suggest some way of working out the individuation of content that can accommodate these problems, then their argument will depend essentially on some specific set of views about content individuation and the way in which sentences are made true in radically different worlds than our own. Part of the strength of the argument from psychophysical harmony is that it does not depend on contentious theses, but as this objection should make clear, the way is not so easy as some perhaps suppose. And as we will see in the next objection, it seems as though there is *no* way of working out the content of our zombie or disharmonious counterparts which is amenable to the argument from psychophysical harmony.

---

[1] Cutter & Crummett apparently draw heavily from Chalmers for their framework, and Chalmers' two-dimensional semantics is generally taken to commit him to the existence of a fairly broad range of narrow contents. It's not clear whether Cutter & Crummett see their argument as relying on two-dimensionalism. But, of course, any controversial view about semantics or the individuation of mental content is going to be a place opponents of this argument (including atheists who find the conclusion unacceptable) can exploit.

[2] This point by itself is compatible with epiphenomenalism, but I do not take epiphenomenalism seriously and so, in presenting this criticism, I am presupposing that the mental is causally efficacious.

[3] Putnam, "The Meaning of 'Meaning'" and Burge, "Individualism and the Mental." Putnam's original thought experiment concerns semantic content; he still supposed in 1975 that Oscar and Twin Oscar are having the same thoughts. He later conceded to Burge that, in fact, both the semantic and representational mental contents of Oscar and Twin Oscar are individuated anti-individualistically, such that Oscar and Twin Oscar don't only mean something different when they say things aloud about "water," they also think different things when they have thoughts which include "water."

[4] This is not some radical minority position; according to the 2020 PhilPapers survey, 58% of philosophers accept or lean toward externalism about mental content, to the 26% who accept or lean toward internalism. I prefer Burge's term "anti-individualism" to "externalism," though in this

case, the point being made about individuation is related to Burge's work on anti-individualism, but is not specifically about anti-individualism; the contents of judgments about phenomenal experience do not constitutively depend on anything "outside" the one who is making the judgment.

[5] Strictly speaking, one *could* make a judgment about water in a world in which there happens to be no water, but there are still oxygen and hydrogen molecules. A scientist might be able to make claims about what would happen if hydrogen and oxygen were combined in such a way as to create H<sub>2</sub>O, and so succeed in referring to it despite never having causally interacted with H<sub>2</sub>O. This point does not help Cutter & Crummett, since they ask us to imagine worlds in which, say, nobody has any experience of redness at all. That would be more like imagining a world in which we make claims about water but neither hydrogen nor oxygen exist at all.

[6] There is an alternative here: Cutter & Crummett agree that in color inversion scenarios, my counterpart might say "blue" when I say "blue," but he might mean "red." This is fine, they say, but cannot account for judgments of similarity. But who is to say that in such a world, *the word "similarity" doesn't also mean something different!*

### Objection #3: Disharmonious Doubles: Liars, Lunatics, or Linguistically Confused?

*By: Lucas Collier & Sebastian Montesinos*

#### *Introduction*

In this section, we draw inspiration from Nigel Thomas' 1998 paper *Zombie Killer*. In that paper, Thomas contends that the conceivability of zombies is made dubious once one tries to interpret a zombie's statement that they are conscious. Thomas argues that any understanding of their report (whether as true, false, or meaningless) is devastating for the conceivability of a zombie world. We believe that Cutter & Crummett also face a very serious analogous challenge in promoting the conceivability of semantic harmony. There is no clear way to interpret our disharmonious doubles' reports about their phenomenal experiences that do not lead to the inconceivability of such worlds or to the collapse of the psychophysical harmony argument.

Take the example of semantic harmony given by Cutter & Crummett where someone (we'll call him Fin) looks upon a ripe red tomato. In a harmonious world *v*, Fin would have the phenomenal experience of seeing red before accurately reporting "I had a phenomenally reddish experience." We will also stipulate that the language of *v* is identical to the language of

the actual world (e.g. what 'reddish experience' means in  $v$  is identical to what 'reddish experience' means in the actual world). Cutter & Crummett describe an alternate world  $w$  that is identical to the harmonious world  $v$  except for that Fin instead has the phenomenal experience of seeing green when he looks upon the red tomato. Nonetheless he reports after the fact "I had a phenomenally reddish experience." In  $w$ , Fin's reports are disharmonious with his experiences. So when he reports "I had a phenomenally reddish experience," how should we understand it?

### *Horn #1: True*

The first horn is that Fin is correct in  $w$  when he gives his report. Obviously, this cannot be because his report of a 'reddish experience' actually corresponds to a phenomenal experience of redness, because as per how we have described  $w$ , Fin has a greenish phenomenal experience. Therefore, if Fin is correct, it must be because in  $w$  'reddish experience' means what 'greenish experience' means in  $v$  (and the actual world). If this is the correct understanding of Fin's report, then this is not a legitimate instance of psychophysical disharmony. An important aspect of semantic disharmony is that we conceive of a world physically-behaviorally identical to a harmonious one except we swap out or eliminate phenomenal experiences. But if there is a linguistic difference between  $v$  and  $w$  then there is a difference in verbal dispositions and therefore physical-behavioral states [1].

Imagine in  $w$  that Fin is presented with a red tomato and a green tomato. Upon looking at the red tomato, he would report that he had a 'greenish experience', and upon looking at the green tomato he would report that he had a 'reddish experience.' In  $v$ , Fin would instead identify the phrase 'reddish experience' with the red tomato and the phrase 'greenish experience' with the green tomato. What this situation shows is that  $w$  is a world distinguished from  $v$  by linguistic, conceptual, and (importantly) behavioral differences in people, not genuine semantic disharmony: Fin's concept of red in  $w$  is his concept of green in  $v$ , and vice-versa, and thus he is disposed to behave differently in  $w$  than in  $v$  when presented with red and green objects. Therefore, while this situation is conceivable, it is not an instance of semantic disharmony and it will not serve the advocate of the psychophysical harmony argument.

### *Horn #2: Lying*

The second horn is that Fin's report in *w* is a lie. Perhaps in *w* Fin understands 'reddish experience' just as we do in the actual world and he understands that his experience wasn't one. Nonetheless, he states "I had a phenomenally reddish experience." Once again, while conceivable, this is not a legitimate instance of semantic disharmony: the physical-behavioral states underlying Fin's lie in *w* are not the same as those in *v* where he sincerely reports his experience. This is because there is a difference in his cognitive processes and behaviors when Fin lies in *w* compared to when he sincerely reports his experiences in *v*. Furthermore, Fin in *w* does understand that his experience was phenomenally greenish not reddish, and thus the apparent semantic disharmony arises only through his actions.

### *Horn #3: Mistaken*

The third horn is when Fin reports having had a greenish phenomenal experience in *w* he is mistaken and sincerely believes every word he says. Fin's cognitive processes and behaviors are perhaps identical to those in *v* where he sincerely reports having had a reddish experience, but in *w* Fin happens to be wrong. The problem with this option is that it removes our warrant for concluding that the actual world is a semantically harmonious world.

Let us start with how exactly Fin comes to his belief and report that he is having a reddish experience. One option is that this belief is inferred through some prior conceptual framework or folk theory that he is using to interpret his experiences. This option is doubly problematic: first, if Fin is inferring that he had a reddish phenomenal experience through some folk theory then his observation is non-direct, meaning that there must be some observation or correlate to that experience that Fin is using to associate his belief with his 'reddish experience'. However, since Fin uses that exact same folk framework in *v*, whatever it is he is using to correlate his belief with his 'reddish experience' (call this *c*) in *v* must be the same in *w* where his experience is different. This implies that *c* is an illusory correlate to reddish experiences, since *c* can imply that one has had a reddish phenomenal experience even when this is not the case, as in *w*. As we shall soon see if it is not already obvious, this will remove our warrant for thinking that our world is harmonious. Second, if it is true that Fin's phenomenal beliefs are non-direct, and inferential from some folk framework that is demonstrably unreliable, this gets us a large part of the way towards an eliminativism that would undermine the psychophysical harmony argument for reasons articulated in the 'concerns' section of this piece.

The second option is that Fin in *v* knows he had a reddish experience non-inferentially. As Thomas says, if this is true then in *v* there must be some cognitive mechanism which leads

Fin from his phenomenal experience to the true belief that he had a reddish experience. But because  $w$  is meant to be physically-behaviorally indistinguishable from  $v$ , this cognitive mechanism too must obtain in  $w$ . That is, in  $w$ , Fin must have the same cognitive mechanisms as he does in  $v$ , but in  $w$  they take him from his greenish phenomenal experience to the false belief that he had a reddish experience. Thomas compares this to a “stuck fuel gauge, that reads **FULL** whether or not there is any gas in the tank” (p. 3). If this is true, then in  $v$  Fin would have no ground to conclude that he was not in a disharmonious world. Fin could try to locate himself in a harmonious world while in  $v$ , but he would be experiencing the exact same cognitive processes and saying and doing the exact same things in  $w$ . Therefore, there is nothing he could identify in his cognitive processes, behaviors, or actions that could break the symmetry between  $v$  and  $w$  to reliably show him that he is in fact in a harmonious world. One might want to insist that he could point to his experience of genuine reddishness in  $v$  to justifiably place himself in a harmonious world. However, in  $w$  Fin would be taking himself to be pointing to exactly the same thing, with the same confidence as he does in  $v$  where he did actually have such an experience. For all Fin knows, he is just as mistaken in a  $v$  as he is in  $w$ .

The actual world is supposedly an analog to  $v$ —a harmonious world. So we are in the same boat as Fin in being unable to discern our harmonious world from a disharmonious counterfeit. Furthermore, if we cannot, through introspection, behaviors, or cognitive processes show that we are not systematically mistaken about our phenomenal states, then by Cutter and Crummett’s lights we should likely conclude we live in a disharmonious world. Crummett and Cutter state:

The overwhelming majority of sets of conceivable psychophysical laws would have produced disharmony, as would the absence of consciousness altogether. (Cutter & Crummett forthcoming, p. 17.)

Thus, *ceteris paribus*, we not only lack reason to think we are in a harmonious world, but also have reason to think our world is disharmonious, as the “overwhelming majority” of possible worlds are [2]. In other words, the prior probability of any given world being disharmonious is radically higher than the prior probability of that world being harmonious. This means that in perhaps establishing the conceivability of disharmonious worlds by taking horn #3, one also undermines a necessary premise of Cutter & Crummett’s argument: that we live in a harmonious world.

Another dilemma arises here for Cutter & Crummett. Consider Fin in a disharmonious world. We can ask whether Fin knows he is in a disharmonious world. If he does not know then, as we just saw, Cutter & Crummett may lose support for the key premise in their argument that we are in a harmonious world. Psychophysical harmony is the datum which stands in need of explanation, for Cutter & Crummett; if we are not sure that it even obtains, then their argument is in trouble. On the other hand, if Fin does know that he is in a disharmonious world, but, *ex hypothesi*, he behaves the same in such a world as in *v*, then that knowledge is epiphenomenal, which is unacceptable for Cutter & Crummett if they wish their argument to have any force against those who reject epiphenomenalism. The conceivability of disharmonious worlds (as with zombie worlds), would then appear to stand or fall with the conceivability of epiphenomenalism. An argument related to this horn of this dilemma will be explored in objection #4.

### *Final Considerations*

The three horns listed thus far are not exhaustive: one can also hold that Fin's report in *w* is meaningless, as Thomas explores in the case of a zombie's report of consciousness. The primary justification for this horn in the case of a zombie's report of consciousness is that zombies have no phenomenal consciousness, and that by some linking premise this casts doubt on their ability to state meaningful sentences. What about in worlds in which Fin has, instead of no phenomenal consciousness at all, a mismatch between phenomenal states and physical states? As noted in the previous objection, one may take Fin in a disharmonious world to be making *meaningful and true* statements in such worlds. So for instance, if Fin judges "red is more similar to orange than to green" in a world in which this judgment would be false if we made that judgment, then perhaps in that world "red," "orange," "green," or "*similar*" (or perhaps all of them) *mean something different such that the judgment comes out true when made by Fin*. But this dissolves the disharmony, so Cutter & Crummett will not want to accept this possibility (though they furnish no argument against it, and it is a natural way of cashing out what is happening in disharmonious worlds).

We have already seen why it is unlikely that Fin's judgments are *false* in disharmonious worlds (and, even if they are false, why Cutter & Crummett may not be able to avail themselves of that possibility in arguing for semantic harmony). So can Cutter & Crummett instead argue that in disharmonious worlds, judgments regarding phenomenal states are *meaningless*, and that it is lucky that we are in a world in which our words and reports of our phenomenal



experiences are meaningful? It's not clear how they could do so. First, Cutter & Crummett would need to give a clear account of what it is for a judgment or report to be meaningful which results in Fin's reports being meaningless and ours being meaningful (and this, to be clear, would add another dialectical weakness to the argument). Suppose they succeed in giving such an account. It is not clear that this helps the argument from psychophysical harmony. Even if Fin's reports are not meaningful in a disharmonious world, they would be meaningful\*, where meaningful\* is what people in that world are talking about when they say things like "my reports are meaningful." People in that world succeed in communicating as well as we do; were we in such a world, we would have as much reason to attribute meaning (or meaning\*) to others as we have in our world. That world's being disharmonious would not add any reason to be skeptical about the meaningfulness (meaningfulness\*) of others' reports. Even if that meaning\* cannot (on the hypothetical account of meaning we are attributing to Cutter & Crummett) be the same as our meaning, it will succeed in playing the relevant roles in that world that meaning plays in ours. And it's not as though we're lucky that our words are meaningful rather than meaningful\*. This is why Cutter & Crummett need Fin in a disharmonious world to be making the *same* judgment as Fin in a harmonious world, but for the judgment in the former world to be not meaningless but false. And we have seen already that Cutter & Crummett have not given us enough reason to think this is coherent.

While we admit the possibility of horns other than the ones we have considered, this section has at least shown that the three most immediate, straightforward interpretations of Fin's report force Cutter & Crummett to choose between the primary datum in their argument or the conceivability of semantic disharmony.

---

[1] Cutter and Crummett acknowledge that a difference in language can account for some of the apparent disharmony in the instances of color inversion, but not the disharmony in others. Even if this is right, in other scenarios the proponent of psychophysical harmony still must choose a way of interpreting judgements/reports—including judgements which compare phenomenal states.

[2] This is important because one could reject the kind of sensitivity conditions that might underlie the skepticism implicit in the argument here: this condition is something akin to 'the fact that our cognitive processes and beliefs would be identical in the disharmonious world undermines our belief that we are not.' In cases that are plausible counterexamples to these kind of sensitivity conditions, for instance, situations where you would believe that you will lose

the lottery even though you happen to be in the world where you win, you would almost certainly want to point to the overwhelming unlikelihood that you happen to be in that world. In this case, no such alternative solution to rejecting this sensitivity condition is present: in fact, it is overwhelmingly unlikely that you are in the disharmonious world given the fact that disharmonious worlds are vastly more prevalent than harmonious ones!

#### Objection #4: Causation and Explanation

*By: Joseph Lawal*

Cutter & Crummett insist that their argument does not presuppose the causal completeness of the physical, or epiphenomenalism. That is, they take their argument to succeed even in generating a need in those who accept that the mental is causally efficacious to explain psychophysical harmony. The extension of the argument is apparently trivial: instead of appealing to the perhaps infinite number of conceivable associations between physical states and causally inefficacious phenomenal states, Cutter & Crummett here appeal to the enormous number of conceivable “ways in which physical states might causally depend on phenomenal states” (p. 26). To illustrate, the authors invoke again an inversion of pleasure and pain:

If interactionism is true, the actual interactionist laws assign something like the following nomological role to pain: a certain brain state X causes pain, and pain in turn causes avoidance behavior. Here we have a case of hedonic harmony. But if the interactionist laws had been different, this nomological role might have been occupied by pleasure (with brain state X causing pleasure, and pleasure causing avoidance behavior) or by some evaluatively neutral state. In that case, we'd have hedonic *d*isharmony. (p. 26)

The causal dependence of avoidance behaviors on pain is still extraordinarily lucky. If the psychophysical laws had been different than they actually are, a physical stimulus which causes pain in the actual world might have caused pleasure while the avoidance behavior stays the same, such that we would have avoided stimuli which actually caused us to experience pleasure. Thus, it looks like the argument is not affected if we drop the assumption of epiphenomenalism; psychophysical harmony still requires explanation.

On closer inspection, this extension of the argument begins to look strained. A major problem for the move to interactionist dualism (or physicalism which rejects epiphenomenalism) is that Cutter & Crummett don't actually seem to take the notion of interaction very seriously; the

authors describe interactionist dualism in a way that entirely neuters it, rendering it nothing more than a kind of epiphenomenalism+. To see why, we need to consider the relationship between causation and explanation. We need not go too far into the weeds on a complicated set of issues to raise problems for Cutter & Crummett.[1] The basic idea we need is that causation and explanation are interrelated and interdependent notions. We cannot attribute causation independently of our ordinary practices of explanation. I take this to mean that when a suggestion is made that phenomenal state X causes physical state Y, we are entitled to ask if it makes any sense to say that X explains Y. If we decide that X cannot *explain* Y, then we can deny that X can be said to have *caused* Y.

Cutter & Crummett describe an epistemically possible world (a world we can imagine or conceive) in which stubbing your toe causes you to experience pleasure, and that pleasure in turn causes you to avoid stubbing your toe in the future. But is it right to say that the experience of pleasure *causes* you to avoid stubbing your toe in the future? Does it sound coherent to say you avoid stubbing your toe *because* it feels good? [2] In other words, would we think that the pleasure is a satisfactory explanation of your avoiding stubbing your toe? It doesn't look like the right kind of thing to explain such behavior. The attribution of a cause here is made independently of our ordinary practices of explanation, which raises doubts as to whether it's a coherent attribution of causation. Causation here is functioning as a kind of black box – there is a phenomenal input and a physical output, and we are told that in the intervening space there exists some mysterious relationship we can call “causal.” There is an explanatory gap between the phenomenal state of pleasure and avoidance behaviors, and between the phenomenal state of pain and seeking-out behaviors. I'm not insisting that Cutter & Crummett have a solution to the mind-body problem or a worked out account of causation, but I *am* asking them to be responsible to our actual practices of explanation and attribution of causal relationships. As it stands, their description of disharmonious worlds does not meet that requirement, and so no one should feel themselves to be under any pressure to call such worlds “conceivable.”

Part of the problem here, I suspect, is the low bar suggested by terms such as “imagine” and “conceive.” *Surely it's obvious that we can imagine pleasure causing avoidance behavior!* we are told, and those who raise doubts about the conceivability of such scenarios are dismissed as a radical minority. To discourage falling into the trap of agreeing too readily that this is all perfectly intuitive, I suggest thinking in terms of what we can describe, or how we would describe a scenario. Would we describe a scenario in which there is a pleasure state and a subsequent set of avoidance behaviors as one in which the pleasure *explains and causes* the

avoidance behaviors? It is not at all clear to me that we would, should, or really can describe this scenario in this way; merely stipulating that there is a causal relationship does not do anything to distinguish this scenario from an epiphenomenalist counterpart. *Under-describing* the scenario makes it much easier to grant conceivability, but one needn't have strong physicalist leanings to think the way the authors describe these worlds is not apt. There is a serious danger, in the description Cutter & Crummett give us, of falling into a Humean view of causation.

I think we can put this point in a way that makes the problem clearer. Cutter & Crummett, in response to the considerations raised by Mørch and Langsam, admit that "it is at least difficult to conceive of phenomenal states having radically different causal powers with respect to some of their purely mental effects" (p. 28). If this is right, then a phenomenal pain state might (epistemically) necessarily be accompanied by, say, a willing on the part of the subject to avoid whatever stimulus caused the pain. In a hedonically disharmonious scenario, then, physical state X causes phenomenal state Y, a painful phenomenal state. Y causes other mental states – perhaps the verbalized thought "That hurts!" and a desire to avoid the pain-inducing stimulus. Nevertheless, because this is an inverted case, Y also "causes" bodily reactions, including speaking aloud "That feels good!" and behaviors which an observer might describe as appearing to seek out the pain-inducing stimulus.[3] But is it not strange (at best) to describe a case in which all of your mental life is systematically disjointed from your actions as one in which mental events cause and explain the associated physical events? This description also seems to entirely bypass the causal efficacy of non-phenomenal mental states. We talk of desires, intentions, thoughts, etc. causing actions as often (perhaps more often) than we talk of phenomenal states causing actions, but Cutter & Crummett draw a straight ("causal") line from phenomenal states to physical outputs.

Perhaps, despite the apparent partial concession to Mørch and Langsam, Cutter & Crummett wish us to imagine a world in which our phenomenal states are changed but all the rest of our mental life stays the same. This certainly is impossible; if, in this world, I experience pleasure and think "I like *that* feeling," that particular thought *cannot* be the same thought in a world in which the phenomenal state on which it constitutively depends is different. Changes to our phenomenal states will necessarily change other kinds of mental states. But even setting the issue of the individuation of intentional mental states aside, this move does not avoid the objection I am raising. Interactionist dualists want to say that mental states are causally efficacious with respect *both* to the physical *and* the mental. A world in which our phenomenal

states are different but our mental and physical states otherwise remain the same is a world in which phenomenal states are completely detached, epiphenomenal with respect to both the physical and the mental realms.

This objection to Cutter & Crummett is related to the objection they consider by Mørch and Langsam. They differ, however, in important respects. This argument emphasizes the relationship between causation and explanation, and so provides a different avenue into making an argument regarding causation. It also does not presuppose the controversial notion of synthetic a priori, as Mørch's argument does, nor even a robust notion of a priority.[4] It can also dispense with a strong reliance on conceptual analysis and even with the supposition that there is some fixed set of epistemically possible worlds.

The foregoing was an argument focusing on the nature of causation simpliciter. We can extend the argument even further by involving the notion of agency.

The argument from psychophysical harmony ignores the complexity of our mental lives, reducing as it does the relationship between the mental and the physical to a simple causal chain from the physical to the phenomenal and back to the physical. The argument loses all plausibility at this stage once we start to complicate this picture even slightly. The problem is that what is thought of as mere disharmony (say, hedonic disharmony) is really a systematic mismatch of intention and action. In other words, the following objection is supposed to raise the question of what the sense is in which you are an *agent* in a hedonically disharmonious world.

Let's say your spouse rubs your shoulders as you sit at your desk. In the inverted scenario, this is agony, but your body relaxes and you say aloud "that's very nice, thank you." That is not what you had intended to say; it was not nice, and you are not appreciative, since the massage was painful. You are determined to avoid massages in the future because it was so unpleasant. On another occasion, your spouse offers to massage your shoulders again, and you internally say "no, that was extremely painful," but to your horror, you find yourself saying aloud with apparent enthusiasm "yes please!" How is this different from the epiphenomenalist version of this story? You're a passive observer in your own life, totally incapable of making decisions in any meaningful sense with respect to the disharmonious phenomenal states. The only difference is that we're adding as an afterthought: "the phenomenal states are "causing" the behaviors." In this world, you're a passive observer to your physical actions, and yet somehow it is supposed to still be a world in which you cause your actions.

This problem generalizes beyond hedonic inversion. In a world in which all physical stimuli are associated with a phenomenal state like TV static, I will wonder why I sometimes say “What a beautiful dog you have!” and other times say “What an awful smell!” even though my experience never changes. In fact, it looks as though in such a world, I would not know why I was using these words - the behavior would be literally inexplicable. But if, as I have suggested, explanation and causation are closely tied together, we have good reason to think that in describing something *inexplicable*, we are describing something which, in the relevant sense, is *not caused* by the kinds of things we expect it to be caused by (i.e., mental states which relate in the right ways to my stating “What an awful smell!”, etc.).

We can say “it’s lucky that our attempts at actions – certain mental events – don’t just have totally random effects. My picking up a cup could have been *caused* by my intending to stand up instead.” But it’s not clear that this is a coherent thing to say; it just looks like epiphenomenalism. The problem for Cutter & Crummett is that there simply is *not* a simple function from external stimulus to phenomenal state to physical output. Phenomenal states are only part of a complex set of interactions between mind and world, and as soon as we recognize the complex and systematic interrelation between phenomenal states, mental states and events and physical states and events, it becomes clear that the picture cannot be so simple as Cutter & Crummett allege.

The point here is that for anyone who denies epiphenomenalism, it is at best questionable whether the core of the argument from psychophysical harmony has any force at all. The argument depends on my being able to conceive of pain causing me to smile, etc., but the mental’s being causally efficacious seems to undermine the conceivability of such a state of affairs. My mental states causing my physical behaviors, in many instances, involves my status as an *agent*. Cutter & Crummett can’t respond here by saying it’s only in virtue of the harmony God instituted that I am an agent at all, because the problem raised poses a prior problem for their setting up the argument from psychophysical harmony in the first place. My handle on what psychophysical causation looks like depends in part on my understanding of what it is like to be an agent.

---

[1] For a brief account of the relationship between causation and explanation which is amenable to this objection, see the first afterword in Putnam's *The Threefold Cord*.

[2] Barring, of course, unusual circumstances, as in which someone is taught to feel guilty whenever she experiences pleasure, and so avoids pleasure. The avoidance in this case is, in a sense, explained in part by the pleasure, but the guilt is at least as important in the explanation. This is not the kind of case Cutter & Crummett have in mind in any case.

[3] Cutter & Crummett would also have us believe that the spoken sentence "That feels good!" in the inverted scenario means the same thing as what we mean in this world when we say "That feels good!"

[4] Putnam himself, who advocates this account of the relationship between causation and explanation, argued against various notions of a priority throughout his entire career. See especially "Two Dogmas Revisited" and "Rethinking Mathematical Necessity." The key difference between an approach like Mørch's and the one I suggest here inspired by Putnam is that Mørch's approach implies or entails the existence of epistemic necessities or rationally unrevisable statements; Putnam developed his notion of "contextual apriority" specifically in order to reject the existence of such statements.

### Objection #5: Understated Evidence

*By: Sebastian Montesinos & Benjamin (Truth Teller)*

#### *Introduction*

Another objection to the argument from psychophysical harmony is that it runs afoul of the fallacy of understated evidence. This is a common problem with theistic arguments identified by Paul Draper, and it occurs when one identifies a general fact G that fits better with theism than naturalism, while ignoring more specific facts S that fit better with naturalism than theism [1]. In this case, the theist points out that our general harmonious laws fit better with theism than naturalism, but ignores more specific ways in which we face disharmony that fit better with naturalism than theism. We will begin by identifying the many instances of disharmony that pervade our experiences, and argue that the level of disharmony in our world is far greater and more serious than the proponents of this argument have supposed, and that this disharmony is

especially damaging for perfect being theism. Next, we will demonstrate that Crummett & Cutter's general response to this seriously underestimates the force of this objection, and how they have much more work to do to overcome it.

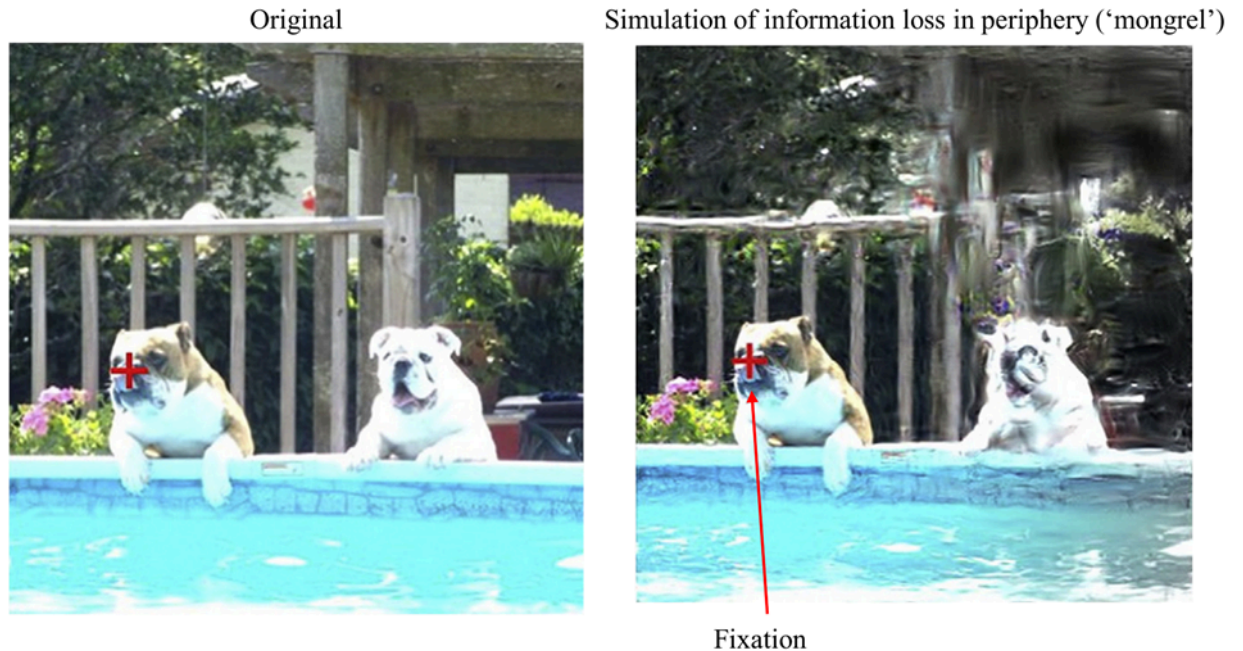
### *Our World is Much More Disharmonious than You Think It Is*

First, human memory is highly flawed and prone to error. Memories and even similar words or phrases are stored in adjacent regions of the brain, and when we recall these things we reconsolidate that information by reactivating the neurons involved in that memory trace, often resulting in us combining disparate pieces of information together, and so confuse details. As an example, in one study participants were more likely to incorrectly pick Alexander Hamilton as one of the former Presidents than some of the actual former Presidents, because the same part of the brain which encodes information about former Presidents also often encodes information about Hamilton [2]. In fact, the mainstream model of memory in cognitive psychology suggests that every time we pull something from our memory that memory becomes temporarily 'fragile' and prone to alteration, and when we restore it we restore its altered clone rather than the original memory. Other psychological research has shown that we can very easily embellish memories and even generate entirely false, elaborate memories, which is one of the reasons eyewitness testimony is so problematic [3-5]. It is not uncommon that people have vividly remembered traumatic events from their childhood such as sexual abuse which never happened often leading to disastrous outcomes, and there are even cases of shared false memories such as the well-known phenomenon of the mandela effect.

Second, humans are prone to addiction to certain drugs, which often have disastrous effects on the human psyche. Surely, there is an epistemically possible world where the human mind is either resistant to bad adverse effects of addictive drugs, or where humans are built with dispositions to avoid harmful drugs and no tendency for addiction towards them. Then there is the human mind's susceptibility to dementia, alzheimer's, cancers, seizures, and other diseases/disorders which severely impair motor function, memory, decision-making, perception, and even personality from damages to certain areas such as the frontal lobe and the cerebrum. If you keep our psychophysical mappings otherwise the same, but take those diseases and other conditions out, it seems to be an epistemically possible mapping much more harmonious than ours.



Third, even our immediate perceptions of the world are riddled with disharmony, far more than most naive observers would expect. Everyone will already be familiar with the existence of hallucinations, and the many visual illusions that people are prone to. What people are likely less familiar with is the extent to which our commonsense beliefs about our perceptual processes are mistaken. We seem to perceive a world full of color across our visual horizon, and to have a detailed-rich picture of the world around us. Yet, research on phenomena such as inattentional blindness, change blindness, and crowding has shown that our visual world is far more limited than we think it is [6-10]. We frequently miss unexpected, obvious objects, we are extremely poor at detecting changes to scenes, and the way our periphery processes information is fundamentally limited. Specifically, it appears that we really just capture the general, rough 'gist' of the statistics of the information in our periphery, without the rich detail [11-12]. To see just how shockingly poor our beliefs about perception are, [look at this video](#) showing the stimuli from an experiment examining some of these limitations. In that experiment, participants viewed a series of scenes within virtual reality, and the periphery of their vision was slowly desaturated (the color was taken out) without their knowledge [13]. The researchers were interested in seeing how much of the world they would need to change before the participants noticed. Shockingly, even in the most extreme conditions, in which all the color in people's visual field was removed except for a tiny dot at the fovea of the eye, a very significant percentage of people still failed to notice these changes. While this experiment focused on color in the periphery, other experiments have shown that our periphery does not represent fine detail, but instead summarizes the average statistical gist of that information, making peripheral vision much more accurately represented by the image on the right below, rather than the image on the left:



What all of this suggests is that our perceptual beliefs are out of line with the actual way in which our perception represents the world. Massively out of line, in fact. Thus research in the cognitive sciences suggests that our everyday perceptual beliefs and experiences are far more disharmonious than most would expect.

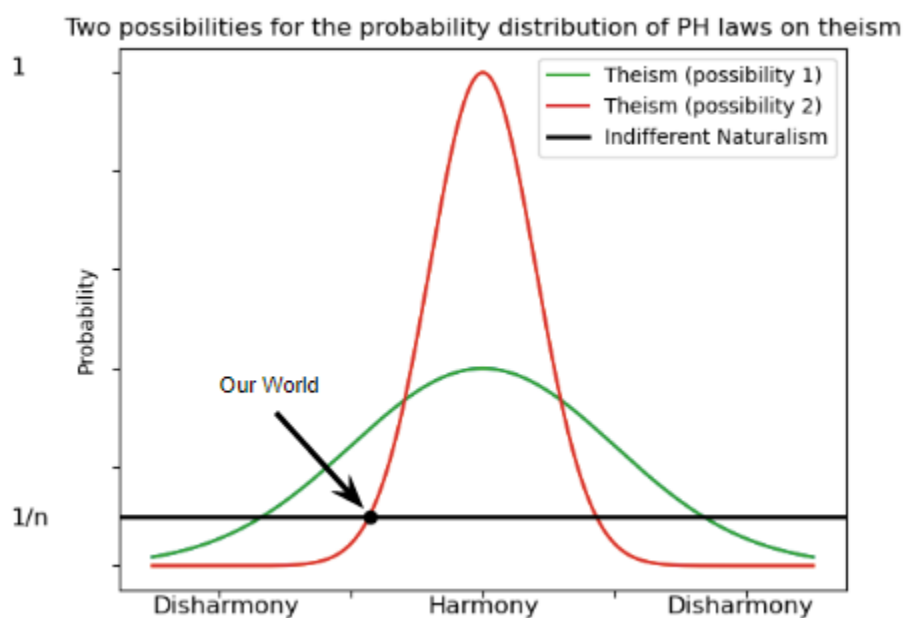
So far, we have discussed ways in which epistemic & semantic harmony run afoul of the fallacy of understated evidence. Let us now turn to normative harmony. Cutter & Crummett rely on the supposition that it is at-least epistemically conceivable that morally salient phenomenal states such as pleasure & pain could be mapped onto different functional profiles. For instance, perhaps pain could be linked to a functional state which disposes us to seek it out, and pleasure to a functional state which disposes us to avoid it. However, we contend that if Cutter & Crummett are right that the phenomenal states of pain & pleasure can be conceivably detached from their functional profiles, then it is also conceivable that we could be disposed to avoid illness & injury, just as we are in the actual world, without those physical states being mapped to painful, or at-least as horrifically painful sensations. Since Cutter & Crummett agree that pain is intrinsically morally bad such that we necessarily have reason to avoid it, surely they must agree that a mapping of physical states to non-pain, more morally tolerable sensations is all-other-things-equal better than a mapping which includes pain, and indeed, particularly horrific instances of it, and yet the latter is exactly the world we find ourselves in. So, that physical states are linked to intrinsically morally bad phenomenal states is a fact about our

psychophysical laws, which must count against theism, since theism generates a predictive expectation towards psychophysical laws which are morally good, and thus a disexpectation for those laws which are morally bad. This point can be further strengthened by the fact that there are many instances of pain which are clearly biologically gratuitous, and do not seem to serve any desirable purpose whatsoever, for instance burning to death does not need to be as agonizing as it is to serve our biological goals, or, it would seem, moral goals either. Furthermore, there are cases where our being disposed to seek out pleasure is bad for us. Pleasure plays a role in addictive behaviors, which very often leads to greatly detrimental long-term outcomes, especially addiction for drugs, alcohol, unhealthy foods, gambling etc. And there are cases where our being disposed to avoiding pain is also bad for us, e.g needles, surgical operations, alcohol on wounds etc.

In addition to the specific ways in which the data of normative and semantic harmony are understated since there are many ways in which our particular psychophysical mappings seem morally undesirable/could have been improved. There are also many, many general conceivably entirely different ways our distribution of psychophysical laws could have been completely different which either seem better or just as good as the actual world. Take the world where all our psychophysical laws are automatically aligned with what an ideally rational version of ourselves would desire, or the world where we can rationally set up for ourselves at any given moment what our psychophysical profile is, or the world where our physical states are all lined up with the experience of a dream-like reality simulating the best life we never could have imagined we wanted, or a world where we can only customize our visual color spectrum to whatever we want, or worlds where our color qualia is automatically set but completely different from ours but such that it makes no practical difference and we can navigate reality just as well, or worlds where our auditory spectrum is completely different but yet we can still navigate and internally comprehend reality just as well. The very same motivations for taking semantic and normative harmony to be conceivable will be reasons to think these alternative distributions are also conceivable. And there seem to be uncountably many worlds like this where the distribution of psychophysical mappings are either better or incommensurable to our actual distribution of psychophysical mappings, thus making it extremely unclear that theism predicts the actual distribution in particular.

*Why Cutter & Crummett's Response to Understated Evidence is Insufficient*

At this juncture, a friend of the psychophysical harmony argument might retort that while theism does not predict our particular distribution of psychophysical mappings, theism predicts those set of distributions which are orderly enough, coherent enough, normatively good enough, and our psychophysical distribution falls in the 'good enough' range, whereas naturalism (or indifference) is so bad at antecedently predicting orderly, coherent and morally good psychophysical laws, that theism still gets the advantage. In fact, [this is exactly what Crummett said in response to this objection when Emerson Green brought it up in this interview](#). However, this is not an adequate response to the objection from understated evidence. To consider why, consider the following possible hypothetical probability distributions for harmonious laws on theism and naturalism:



If, on theism, we should predict an epistemic probability distribution that is strongly weighted towards *the most* harmonious distributions, and on which you have an immediate, precipitous fall in the probability once you approach any mix of harmony and disharmony that contains any non-insignificant amount of disharmony, then theism predicts the data *just as badly or worse as naturalism* ("Possibility 2" on the above figure). On the other hand, if theism implies a fairly inclusive antecedent probability distribution on which worlds that are mostly harmonious but contain significant disharmony are still highly probable, then theism does do better than 'indifferent naturalism' ("Possibility 1" on the above figure). What this implies is that for theism to have explanatory content with respect to this data, *Crummett owes us a principle about what to*

*expect on theism on which we can antecedently expect that the peculiar mix of harmony and disharmony we observe is not very unlikely.* Cutter & Crummett simply do not do this in their paper. Pointing out that harmony is good and that God is perfectly good is, if anything, reason to expect only the most or close to the most harmonious distribution on theism. Additionally, there are a number of plausible principles about what to expect on theism that would lead us to expect this. Thus, at minimum, Crummett owes us some kind of principle on what to expect on theism, and a justification that the specific mix of psychophysical laws we observe fit into this principle.

As an analogy, consider that there is a factory that produces spherical-like objects, and there are two robots that make these objects. The first robot, 'randobot', outputs spherical-like forms where the extent to which they trend toward perfect sphericity is random. They can be lumpy, imperfect, bizarre, etc. The other robot, 'sphereobot' is a robot that is tuned to make the most smooth, perfect spheres possible with no blemishes. Additionally, it has all the materials in the factory it needs to ensure that every sphere comes out as spherical as possible. Now, suppose a sphere comes out that is pretty good looking and trends towards sphericity, but has a number of problems: it has lumps and distortions in a number of places. Which robot should we infer made the sphere? Obviously, we have no good evidence that 'sphereobot' made this sphere and not 'randobot', if anything, we should infer it was randobot.

One principle vis-à-vis God's actions that would certainly doom the psychophysical harmony argument would be a strong, 'best possible world principle'. This principle may be something like: If God wills X, then X is the best possible state of affairs among the range of logically possible options which could have been actualized. It is not enough that God wills states of affairs that are "good enough", rather we should expect, and indeed it is entailed, that God wills the best state of affairs. The principle is plausible given the theoretical content of theism, which posits a maximally good, maximally powerful being. If true, this principle entails that those distributions of psychophysical mappings which have any cases of disharmony relative to other logically possible distributions which are more harmonious, have a probability of 0 on theism. Of course, this principle is notoriously controversial, mainly on the basis that there is no best possible world or state of affairs. Whether this response ultimately favors the theist is controversial [14]. However, it is very important to note that the objection we have given above in no way relies on such a strong principle. As long as there is some plausible principle about what to antecedently expect on theism that is strong enough to make the actual distribution of psychophysical mappings we see highly unlikely, the psychophysical harmony argument is undermined. And, remember that the particular distribution of psychophysical laws we see is not

actually nearly as tuned towards harmony as your everyday observer would think, as indicated by the research cited above. There are significant ways the particular distribution of laws we see could have been improved that do not entail that they are perfectly harmonious. Therefore, it is definitely not sufficient for Cutter & Crummett to simply reject a 'best possible world' principle.

To give just one example of a weaker principle that would doom the psychophysical harmony argument, consider the following: "Necessarily, God does not will any state of affairs which entails horrific suffering upon His creation." Once again, this principle looks plausible given theism's content, (God is taken to be all-loving and perfectly good) and unlike the previous principle, does not lead to a worry of there being no best world. Yet, there are many cases in the actual world where our physical states are mapped to horrifically agonizing phenomenal states. Consider, being burned alive, severely beaten and tortured, or severe cases of epidermolysis bullosa. Cutter & Crummett might continue to insist that the above principle is as well, too strong, and that there are plausible theodical reasons why God might allow such suffering. But then it is on Cutter & Crummett to provide such reasons, and to motivate a principle that is plausible given theism's theoretical content, and is such that theism antecedently predicts a broad enough range of psychophysical distributions so as to include ours.

Given how much disharmony we actually observe, and the peculiar ways it manifests, we find it eminently questionable that our distribution of psychophysical laws falls in the "good enough" range all-things-considered on any plausible account of what to expect on theism. However, suppose Crummett & Cutter can actually lend theism explanatory content by maintaining a weak principle that allows theism to predict a broadly "good-enough" range for the psychophysical mappings. Even granting this, the understated evidence objection still successfully shows, at-least, that the psychophysical harmony argument doesn't have nearly as much evidential weight as we would initially think ([In particular as much as Apologetic Squared hypes it up to have](#)). It is not the case that the epistemic probability of our psychophysical laws on theism is high and on naturalism it's astronomically low. It is astronomically low on both, just at best less low on theism to an unknown degree, since while naturalism has no restriction on the psychophysical laws we would antecedently expect, theism restricts the psychophysical laws we would expect to a "good enough" range, which insofar as it includes our particular distribution, as we discussed, it also contains a vast, inscrutable, possibly infinite amount of epistemically possible distributions, some much more antecedently predicted on theism than ours in virtue of being morally preferable, and theism has no resources to predict ours in particular relative to other possibilities within that range. So while, given this concession, psychophysical harmony would be some evidence for theism, it would not be incredibly strong

evidence and it would not at all be clear that the posterior probability will come out in favor of theism.

---

- [1] Draper 2010
- [2] Roediger & DeSoto 2016
- [3] Loftus 1997
- [4] Loftus 2019
- [5] Talarico & Rubin 2003
- [6] Neisser 1979
- [7] Jensen et al. 2011
- [8] Levin & Simons 1997
- [9] Whitney & Levin 2011
- [10] Cohen et al. 2021
- [11] Rosenholtz 2016
- [12] Rosenholtz 2017
- [13] Cohen et al. 2020
- [14] Rowe 1994

### Objection #6: The Revenge Problem

*By: Benjamin (Truth Teller)*

#### *Introduction*

This last objection is known by friends and foes of the psychophysical harmony argument as "the Revenge problem". In brief, the problem is that if it is the case, as the psychophysical harmony argument requires, that our orderly coherent distribution of psychophysical laws has an a priori extraordinarily low epistemic probability, then this same fact would apply the features of God's mind. God's mind has an orderly, coherent, and normatively perfect structure such that He accurately perceives external objects and their properties, He never has any false recollections, His will and creative action perfectly aligns with and indeed entails what will be, all propositions He believes are true etc. While I am not extremely certain that this objection works, I think it is a serious concern. I will defend that there is at least a

non-negligible chance that it works, and if it does, I believe it is catastrophic for the argument. This is because, it would imply, by the proponent of the psychophysical harmony argument's own lights, that theism should be assigned a very low prior probability. So low that it is not clear that the data of psychophysical harmony would give theism a sufficient boost in posterior probability. It seems God's mind, being infinite, would be highly complex, and that there are a great deal of epistemically conceivable ways God's mind could have different 'psycho-divine' (Shorthand for the laws governing God's mind) properties, so that it would be highly improbable for God's mind to have the precise psycho-divine harmony it has. Therefore, the theistic hypothesis would have an extraordinarily low prior indeed.

#### *Counter #1: Metaphysical Necessity*

The theist has responses to this objection. One objection I want to immediately dismiss is that God's psycho-divine attributes are brute metaphysically necessary truths about God's nature, so it is a probability of 1 that God's mind has psycho-divine harmony. Notice that the probabilities we are considering are epistemic, it does not matter that God's mind being disharmonious is metaphysically impossible, what matters is that disharmony in God's mind is epistemically possible, e.g it could have for-all-we-know been the case, and that we do not seem to have reason to privilege God's particular harmonious distribution of psycho-divine states over alternative epistemically possible distributions. Even if the theist could make use of this response, the naturalist could as well present the very same objection to the psychophysical harmony argument; that our psychophysical laws are metaphysically necessary. This is leaving aside the problems with metaphysical necessity in general, which is a shaky and controversial notion.

#### *Counter #2: Simple God*

A better objection, and this is the one Cutter & Crummett are sympathetic to, is that God can be characterized in a way that is both intrinsically simple and entails psycho-divine harmony. For instance, God can be taken to be a perfect being, and this simply entails that God's psycho-divine states will be harmonious, or that God is defined as omnipotent, and omniscient and this entails harmony in His mind. I find this line of reasoning immediately dubious, if it is the case that perfection or omnipotence & omniscience implies this complex set of psycho-divine attributes in the bearer of those properties, then this seems to suggest that a



being who exemplifies these properties is not such a simple one afterall. At this juncture, the disagreement would presumably shift to the intrinsic probability of theism in general, of course the theist may appeal to Swinburne-style or Rasmussen-style considerations in an attempt to show that theism is not so intrinsically improbable. I do not wish to take us far afield, but I will argue that these considerations won't work here, you cannot use perfection, omnipotence or God's other attributes to explain His psycho-divine harmony. The reason is because it seems to get the order of explanation wrong. It cannot be the case that God's perfection explains why there is psycho-divine harmony in God's mind, because it seems that God's being perfect is logically posterior to God's possessing psycho-divine harmony. Another way to put the point is that perfection, like goodness, seems to be a thin axiological property. It cannot be logically prior to thick perfect-making properties, which would presumably include psycho-divine attributes, because it would then seem to be an empty predicate. As an example, suppose I claim that my dog is good, you'd take my claim to be about my dog's behavior or other descriptive features, He doesn't bite, He doesn't bark too often, He is obedient, friendly, loyal, cute etc. But suppose I reply, "No, my dog is not good in virtue of any of these good-dog making properties, my dog's goodness is logically prior to these properties but entails those properties". Then it seems like my claim that my dog is good is devoid of any content that would make it intelligible. The same point appears to be true of God's perfection, for it to have content, it has to be true that God instantiates perfect-making properties logically prior to His having the thin axiological attribute of "perfection", for it is those properties which gives His perfection content. There are also problems if we use omnipotence & omniscience to explain psycho-divine harmony. It seems like God's omnipotence & omniscience is logically posterior to harmony in His mind. Surely, it cannot be the case that His omnipotence is causally responsible for His harmonious faculties. He must exemplify various harmonious psycho-divine properties first in order to do so, He cannot pull Himself up by His metaphysical bootstraps as it were. This is analogous to why I take it, God cannot ground all exemplifiabes. For God to ground exemplifiabes there must be an exemplifiable logically prior to God grounding exemplifiabes, be it something like God's exemplifying the ability to ground exemplifiabes. So too, God must exemplify harmonious psycho-divine faculties logically prior to His omnipotence which requires them. Thus, it is not that omnipotence, and perfection or what have you are simple properties which just entail highly complex psycho-divine harmony, it is rather that a being that is perfect, omnipotent, omniscient etc. must, logically prior, be a being with highly complex psycho-divine laws governing their mind.

### *Counter #3: Simple Laws*

Another possible objection is that the psycho-divine laws governing God's mind are themselves very simple, and thus not intrinsically improbable. They can be characterized as something like, "If God wills something, it is so" "If God believes something it is true " etc. The issue with this is that it seems the same law or set of laws, or propositions could be simple under one description and highly complex under another description. For example, (There are exactly 7 swans) is a simple atomic proposition, but it can also be stated as a massive disjunctive proposition (It is not the case that there are no swans & and it is not the case that there is 1 swan, & it's not the case that there are 2 swans & it's not the case that there are 3 swans and so on for every non-7 number) which is obviously highly complex in the number of ways it can turn out to be false. God's laws could be stated as "God can will anything" or it can be stated as "God can will X, & Y & Z and so on" for some infinite, or astronomically large set of states of affairs. We might think, for the same reason we think that there are numerous epistemically possible psychophysical distributions, that there are countless epistemically possible distributions where God has the power to will 1 or 2 or 3, or such-and-such particular state of affair(s) less. The common theistic response to this will be that distributions which contain no arbitrary limits will be more probable, then distributions without said limits. Again, this will take us into general concerns about the intrinsic probability of theism. I will just note for our purposes that this point can be granted, that there is some weight in favor of distributions of psycho-divine laws with no arbitrary limits, but as long as it is not absurdly high, and is a finite  $n$  degree higher than distributions with arbitrary limits then it's still going to turn out that God's mind being harmonious is extremely improbable since there are just so many, most likely infinite, epistemically possible ways God's mind could be disharmonious.

### *Counter #4: Classical Theism*

The last objection I will look at is the classical theist response, perhaps God isn't a mind, so there is no question of psycho-divine harmony in the first place. I'll just say a couple things about this. For one, the classical theists believe that God has an intellect and a will, and is omniscient, these seem to be properties of a mind, so it doesn't seem intelligible on their view to deny that God is a mind. The classical theists' answer to this is that these are analogical predications of God, He does not actually have a mind in any sense we have a mind. I'm very skeptical of the intelligibility of these claims and I have previously expressed such skepticism on

my blog. What is analogy, if not picking out shared features? If they are analogical predicates, it seems there must be shared features, such that "intellect" "will" and "knowledge" is something that can be meaningfully predicated of both God and humans/other creatures which are part of our background understanding. The problem is that our notion of "intellect" and "will" just is as mental properties, it is constitutive of our understanding of these predicates that it is embedded in our broader concept of the mind and if you take this out, we're no longer talking about the same concept. For two, a more general issue I have with an appeal here, to views wherein God is not a mind, is that it seems what makes theism distinct as an explanatory hypothesis from views such as naturalistic axiarchism which has a tendency to predict valuable states of affairs, is that it is an agential explanation. It appeals to the beliefs, and desires, of an agent, aimed at some end, or state of affairs, as well as a supposition that the agent will act rationally in accordance with these desires and beliefs. But if we take out God's being a mind, then it cannot be the case that theism is an agential explanation, and it is unclear how, if at all, we distinguish theistic explanations of the data from naturalistic axiarchic explanations. If we cannot appeal to the virtues of an agential explanation relative to non-agential explanations, or reasons to think an agent in this particular case better accounts for the data, it does not look like the theist would have any resources at their disposal at all to show their view to be preferable to naturalistic axiarchism.

### *Conclusion*

I believe I have covered, and provided reasonably plausible responses to the most popular objections to the revenge argument, and so I hope this sufficiently motivates the claim that, at least, the revenge problem is a while perhaps not insuperable, potent objection to the psychophysical harmony argument which ought to be taken very seriously.

### Conclusion

Here is how the dialectic stands, as we see it. If the argument from psychophysical harmony is sound, it constitutes reason to accept one of a wide range of positions. Among those positions are several variants of theism - perfect being theism, classical theism, and orthodox Christian theism. In fact, any form of theism which can explain the harmony of our phenomenal

and physical states is compatible with the argument, so severely limited gods fall out of this argument as well as unlimited ones (though of course, there may be more specific problems with one or another view). However, the argument from psychophysical harmony is also an argument for what are appropriately thought of as atheistic alternatives, or at least alternatives which many atheists will find more palatable than perfect being theism, not least because many of these alternatives will not be susceptible to such standard atheistic arguments as the problem of evil and the problem of divine hiddenness. Thus, even if the argument is sound, it leaves an enormous gap which needs to be filled by any *Christian* apologist who intends to use the argument for apologetic purposes.

That's all if the argument is sound. One thing we hope to have highlighted here is that, despite claims to the contrary, the argument from psychophysical harmony rests on a large number of highly contested or contestable premises. In this regard, in fact, it may be at a serious disadvantage when compared to more traditional arguments for the existence of God (like arguments from contingency or fine tuning arguments), many of which purport to get closer to the kind of God Cutter & Crummett want to accept anyway. We pointed out several positions which take the wind out of part or all of the argument from psychophysical harmony. For those with eliminative or reductive leanings, a priori physicalism prevents the argument from gaining any ground at all. Those who want to preserve a more robust mental life can look to Davidsonian and Putnamian methods of denying the epistemic gap. And Kripkean conceivability arguments may allow one to deny the conceivability of the scenarios on which Cutter & Crummett without endorsing any theory of mind at all. Some philosophers will be unperturbed by elements of the argument because of prior commitments to moral anti-realism or eliminative materialism. The argument will have little to say to anyone who already accepts some of these views (it's not as though philosophers only feel drawn to, say, Davidson's anomalous monism in order to avoid accepting psychophysical harmony). And, once more, it may be more rational for some naturalistic atheists to accept one of these views over the conclusion of the argument from psychophysical harmony.

Some, particularly those who do not already accept some of the above views or do not find them plausible, will be unsatisfied with such responses. Many philosophers, after all, accept much of the picture of the mind which motivates the argument from psychophysical harmony. For any in this position, we have provided additional reasons to doubt the soundness of the argument from psychophysical harmony. As we showed in objection 1 above, normative harmony looks to be undermined regardless of whether we take introspection seriously; if we

do, we seem unable to divorce states like pain from their functional profiles. If we don't, then the normative status of such states looks contingent on their functional profile. Objections 2 and 3 target semantic harmony, calling into question the coherence of the description of this kind of harmony Cutter & Crummett provide, and raising serious doubts about whether disharmonious worlds can do the work that they need it to with respect to semantic harmony. Objection 4 shows how narrow a conception of causation Cutter & Crummett need for their argument to work against anyone who rejects epiphenomenalism. If we take seriously our ordinary attributions of causation and our ordinary practice of explanation, we should not think that Cutter & Crummett have succeeded in extending their argument to challenge those who reject epiphenomenalism. Objection 5 shows that disharmony is much more present in our world than most probably suppose it to be, and that Cutter & Crummett have much more work to do to show that theism can actually predict this unique pattern of harmony and disharmony better than naturalism—merely stating that on theism we expect things that are 'good enough' won't cut it. And objection 6 points out that if the argument from psychophysical harmony goes through, it appears to pose sharp problems for the theist, in which case the argument from psychophysical harmony better supports the "theism-adjacent" hypotheses which many friends of the argument want to reject.

All in all, atheist naturalists should not lose sleep over the argument from psychophysical harmony. It constitutes an interesting and innovative contribution to the ongoing discussion in the philosophy of religion, but has weaknesses which are not difficult to exploit, and a substantial gap between its conclusion and the conclusion which many of its advocates are seeking.

### Works Cited

- Bourget, D. & Chalmers, D. J. (ms) "Philosophers on Philosophy: The PhilPapers 2020 Survey".
- Cohen, M.A., Botch, T.L., & Robertson, C.E. (2020). "The limits of color awareness during active, real-world vision". *Proceedings of the National Academy of Sciences USA*, 117, 13821-13827.
- Cohen, M., Ostrand, C., Frontero, N., & Pham, P.-N. (2021). "Characterizing a snapshot of perceptual experience". *J Exp Psychol Gen.* <https://doi.org/10.31234/osf.io/nm5uc>
- Cutter, Brian & Crummett, Dustin (forthcoming). "Psychophysical Harmony: A New Argument for Theism". *Oxford Studies in Philosophy of Religion*.
- Dennett, Daniel C. (1995). "The unimagined preposterousness of zombies". *Journal of Consciousness Studies* 2 (4):322-26.
- Devitt, Michael (1990). "Transcendentalism about content". *Pacific Philosophical Quarterly* 71 (December):247-63.
- Draper, P. (2010). "Cumulative Cases". In C. Taliaferro, P. Draper, & P. L. Quinn (Eds.), *A companion to Philosophy of Religion* (pp. 414–424). essay, Wiley-Blackwell.
- Jensen, M. S., Yao, R., Street, W. N., & Simons, D. J. (2011). "Change blindness and inattentional blindness". *WIREs Cognitive Science*, 2(5), 529–546.  
<https://doi.org/10.1002/wcs.130>
- Kripke, Saul A. (1980). *Naming and Necessity: Lectures Given to the Princeton University Philosophy Colloquium*. Cambridge, MA: Harvard University Press.
- Levin, D. T., & Simons, D. J. (1997). "Failure to detect changes to attended objects in motion pictures". *Psychonomic Bulletin & Review*, 4(4), 501–506.  
<https://doi.org/10.3758/bf03214339>
- Loftus, E. F.  
(1997). "Creating false memories". *Scientific American*, 277(3), 70–75.  
<https://doi.org/10.1038/scientificamerican0997-70>  
(2019). "Eyewitness testimony". *Applied Cognitive Psychology*.  
<https://doi.org/10.1002/acp.3542>
- Neisser, U. (1979). "The Control of Information Pickup in Selective Looking". In E. J. Gibson & A. D. Pick (Eds.), *Perception and its development: A tribute to Eleanor J. Gibson* (pp. 201–219). essay, L. Erlbaum Associates.
- Putnam, Hilary (1999). *The Threefold Cord: Mind, Body, and World*. Columbia University Press.
- Roediger, H. L., & DeSoto, K. A. (2016). "Recognizing the presidents". *Psychological Science*, 27(5), 644–650. <https://doi.org/10.1177/0956797616631113>

Rosenholtz, R.

(2016). "Capabilities and Limitations of Peripheral Vision". *Annual Review of Vision Science*, 2(1), 437–457. <https://doi.org/10.1146/annurev-vision-082114-035733>

(2017). "Summary statistic encoding plus limits on decision complexity underlie the richness of visual perception as well as its quirky failures". *Journal of Vision*, 17(10), 202. <https://doi.org/10.1167/17.10.202>

Rowe, William L. (1994). The Problem of No Best World. *Faith and Philosophy* 11 (2):269-271.

Thomas, Nigel J. T. (1998). "Zombie killer". In Stuart R. Hameroff, Alfred W. Kaszniak & A. C. Scott (eds.), *Toward a Science of Consciousness II*. MIT Press.

Talarico, J. M., & Rubin, D. C. (2003). "Confidence, not consistency, characterizes flashbulb memories". *Psychological Science*, 14(5), 455–461.

<https://doi.org/10.1111/1467-9280.02453>

Whitney, D., & Levin, D. M. (2011). "Visual crowding: a fundamental limit on conscious perception and object recognition". *Trends in Cognitive Sciences*. 15(4), 160–168.

<https://doi.org/10.1016/j.tics.2011.02.005>

Yablo, Stephen (2000). "Textbook Kripkeanism and the Open Texture of Concepts". *Pacific Philosophical Quarterly* 81 (1):98–122.