

# On the alleged explanatory impotence/conceptual vacuity of substance dualism

James Moreland 

Department of Philosophy, Biola University,  
13800 Biola Ave., La Mirada, California  
90639, USA

## Correspondence

James Moreland, Department of  
Philosophy, Biola University, 13800 Biola  
Ave., La Mirada, CA 90639, USA.  
Email: [jp.moreland@biola.edu](mailto:jp.moreland@biola.edu)

## Abstract

In the last decade, there has been a notable upsurge in property (PD) and generic substance dualism (SD). By SD I mean the view that there is a spiritual substantial soul that is different from but variously related to its body. SD includes Cartesian, certain forms of late Medieval hylomorphic (e.g., Aquinas'), and Haskerian emergent SD. Nevertheless, some form of physicalism remains the majority view in philosophy of mind. Several fairly standard objections have been raised against SD, and SDists have been preoccupied with these objections. As a result, a potent objection has not been given the attention it deserves. The purpose of this article is to fill that lacuna by raising the visibility of this objection and providing plausible responses to it. First, I shall clarify the explanatory-impotence challenge and, second, provide defeaters against it. Note that several of the issues to follow apply both to PD and SD. I will focus mostly on SD, but sometimes PD is a part of the discussion in that certain arguments for and against PD indirectly affect SD. When it is important for me to note that either SD or PD is specifically being addressed, I will notify the reader.

## KEYWORDS

chimeral colours, peripheral anomalies, substance dualism, conceptual vacuity of, substance dualism, empirical inadequacy of, theory carriers

## 1 | PHYSICALISM AND THE EXPLANATORY IMPOTENCE OF SD

### 1.1 | The objection

In an important critique of SD filled with a number of objections, Ravenscroft (2018, 267–282) asserts the following:

In this subsection, I will advance what I take to be the most serious objection to SD—the observation that physicalist theories of the mind have vastly more explanatory power than SD [italics mine].<sup>1</sup>

Along similar lines, Bayne (2018, pp. 225–228) complains that with respect to a wide range of problems regarding consciousness, while physicalism has a solid track record of explaining such problems and offers the prospect of accounting for further difficulties, “[s]ubstance dualists offer no answer to these questions...” Finally, Murphy (1998, p. 18) opines that.

science has provided a massive amount of evidence suggesting that we need not postulate the existence of an entity such as a soul or mind in order to explain life and consciousness.

This evidence consists of the fact that “biology, neuroscience, and cognitive science have provided accounts of the dependence on physical processes of *specific* faculties once attributed to the soul” (Murphy 1998, 17).<sup>2</sup> Elsewhere, she (2006, p. 56) claims:

My argument in brief is this: all of the human capacities once attributed to the mind or soul are now being fruitfully studied as brain processes – or, more accurately, I should say, processes involving the brain, the rest of the nervous system and other bodily systems, all interacting with the socio-cultural world.

Murphy acknowledges (Murphy 2006, p. 112) that.

dualism cannot be *proven* false – a dualist can always appeal to correlations or functional relations between soul and brain/body – but advances in science make it a view with little justification.

It is important to note that in these and other objections, there is a tight connection between the nature and appearance of consciousness (typically, an issue regarding PDism) on the one hand and the explanatory impotence of SD in explaining consciousness on the other.

To clarify this connection, while debates about PD are different than those about the possessor and unifier of consciousness (SD),<sup>3</sup> the issue before us explicitly brings these together because as we will see later, some physicalists argue for the explanatory impotence of SD either because consciousness is physical and, thus, there is no

<sup>1</sup>Especially 276–278. The quote is from 276.

<sup>2</sup>cf. 13, 27.

<sup>3</sup>An anonymous referee rightly pointed out to me that debates about the non-physical nature of the possessor and unifier of consciousness are not necessarily the same as debates about SD. See Nida-Rumelin (2010). I agree, but for my purposes, I do not think it is relevant. In context, I am merely claiming that debates about the possessor and unifier of consciousness are different than those about PD. The debate about the former would include those involving the adequacy of SD vis a vis. either alternatives to SD such as Nida-Rumelin's view or physicalist attempts to explain or provide reductive analyses of the nature of the possessor of consciousness and the sort of unity that characterizes it.

positive characterization of a soul (it is inadequately characterized by the *via negativa* as immaterial) or because even if PD is true, postulating a soul as a causal explanation of the obtaining of conscious states is of no help at all (we have no clearer idea of how a soul causes conscious states than, say, the brain.)

SD is either a view about the conscious essence of the soul (as SDists claim) or that which causes and explains consciousness (as many critics hold).<sup>4</sup> If the former, then physicalist theories of consciousness as proffered by Paul Churchland (see below) leave the intrinsic nature of a soul conceptually opaque and, thus, count against its existence. If one asserts the existence of glumstock and there is no concept of what glumstock is supposed to be, the existence claim is unintelligible and hard to take seriously. Since SDists employ (at least) the actual and dispositional properties of consciousness as depicted by PDists to provide a positive account of a soul's essence, by taking these to be physical properties, and rejecting the adequacy of a purely negative characterization of the soul, "soul" becomes like "glumstock." If the latter, the same result follows since now the soul is not only explanatorily irrelevant—and, thus, there is no sufficient reason to believe in it—but also the concept of a soul turns out to be a vacuous, unintelligible sort of stuff bereft of any positive characterization. Consequently, in what follows, I will sometimes address physicalist critiques of PD which, in the two ways just noted, support the empirical explanatory inadequacy of SD, and I will sometimes address SD when the relevant arguments directly involve problematic features of the soul itself or its explanatory impotence.

## 1.2 | Rejoinders to the objection against SD

What should one say about this line of reasoning? I offer five responses. First, many substance dualists do not believe in a soul primarily because it is a theoretical postulate with superior explanatory power. Rather, they take the soul to be something of which people are directly aware. As Hopp (2011, p. 4) succinctly puts it in regard to the phenomenology of consciousness: "For many...approaches to the mind, conscious experiences are theoretical postulates postulated to explain certain observable data. For the phenomenologist, they are the data."

For many SDs, Hopp's point applies with equal force to the soul. I take Lycan's (2018, p. 27) more general statement to be canonical: "Cartesian minds are not explanatory posits at all, much less posited to explain physical facts. They are known from the inside and there is nothing queer or obscure about that."

Thus, belief in a substantial soul is properly basic and allegedly grounded in self-awareness. By "properly basic" I mean a belief that is rightly taken to be (defeasibly) justified, e.g., by experience, without needing to be justified by other beliefs. The point is not that dualists are right about this. Rather, given this dualist approach, the point is that advances in our knowledge of mental/physical dependencies are simply beside the point for such SDists. Dualist entities—conscious states and the soul—are not embraced primarily due to their theoretical or explanatory superiority. They are entities of which one seems to have direct acquaintance. And the further debate about which approach (the direct awareness or the theoretical postulate approach) is the fundamental one for defending SD is not something for which advances in scientific knowledge are relevant.

Second, in those cases where a soul is postulated as the best explanation for a range of purported facts, typically, those facts are distinctively philosophical and seldom the scientific ones Murphy mentions. Arguments from the unity of consciousness, the most plausible view of that which is essentially characterized by and the possessor of consciousness, the metaphysical possibility of disembodied survival or body switches, the best view of an agent to support libertarian agent-causation, the metaphysical implications from the use of the indexical 'I', and the special sort of diachronic and synchronic unity of human persons are typical of arguments offered by SDists. Accordingly, the fact that SD does not explain the sort of facts treated by empirical neuroscientific theory is fundamentally beside the point. SD is a philosophical position not primarily proffered to explain those facts.

<sup>4</sup>For examples of this claim, see Rickabaugh and Moreland (2023), chapter ten.

These two factors—direct awareness of the soul and the soul as the best explanation for a range of distinctively philosophical issues—are *the* fundamental epistemic grounds for SD.

Murphy's position about the scientific nature of SD is not benign. She misrepresents the debate between SDists and physicalists since her depiction of it assumes a presupposition that she neither defends and that SDists reject (the fundamental importance of empirical adequacy for SDism's credibility). Interestingly, when one examines specific examples of physicalist critiques of and alternatives to PD/SD, it is the philosophical aspects of the debate that do the heavy lifting. To cite a curious example of this, in a widely known book, Churchland (2013) explicitly rejects so-called First Philosophy, claims that the empirical neuroscientific issues are the fundamental ones in the dispute, and adopts reductive physicalism for sensations and eliminativism for propositional attitudes. However, the first half of the book is where Churchland presents his case against PD and for his physicalist views. He concludes that this case renders irreducible consciousness and the soul obsolete and discredited. The last half of the book discusses empirical issues, and they play virtually no role in building his case. That is played out in the book's first half which is entirely philosophical.

I am not suggesting that Churchland does not argue elsewhere that empirical factors favour physicalism. I will address this below. For now, I merely point out that, despite Churchland's claims to the contrary, his actual argumentation against PD (and, consequently, against SD) and for his physicalist positions are philosophical. I suggest that examples like this are widespread in the literature, and they show that dualism is fundamentally recognized to be a philosophical position. Empirical factors may, indeed, play a role, but that role is quite secondary, given the nature of dualism as it is presented by its advocates. To claim otherwise comes close to being a strawman fallacy. The elevation of empirical issues by Murphy and others gives those issues more epistemic importance than they deserve.<sup>5</sup>

Third, some dualists (Manzotti and Moderato 2014; Antonietti 2008) supplement their philosophical arguments for SD, by secondarily bolstering their view in light of neuroscientific data. Even non-dualists (Bennett and Hacker 2003; Rockwell 2005; Uttal 2001, 2004) recognize that neuroscientific methods are based on dualism's conceptual framework, as they take first-person reports to be privately accessible to and epistemically privileged for the reporting subject for establishing mental/physical correlations and SDists argue that this private, privileged access in one way or another provides sufficient grounds for SD.<sup>6</sup> More on these last points below.

Fourth, others argue that neuroscience fails to support physicalism over dualism.<sup>7</sup> Space considerations prevent me from going into detail about this issue. I simply note it is employed by several SDists. Thus, if one is to defend the conflict between neuroscience and dualism, one must offer plausible replies to these last two lines of argument.

Finally, the discovery of “the dependence on physical processes of *specific* faculties once attributed to the soul” does not provide sufficient grounds for attributing those faculties to the brain rather than to the soul. The first thing to note is that there is an important distinction between describing the nature, proper categorization, and possessor of a capacity versus explaining what conditions are necessary for its actualization. To see this, it is important to get clear on the use of “faculty” as the term has been historically used in discussions of substances in general and the soul in particular.<sup>8</sup> Roughly, a faculty of some particular substance is a natural grouping of resembling capacities or dispositions possessed by that thing. For example, the various capacities to hear sounds would constitute a person's auditory faculty.

<sup>5</sup>I agree with E. J. Lowe that the principle of causal closure is metaphysical and not empirical. See Lowe (2008).

<sup>6</sup>For arguments of this sort, see Rickabaugh and Moreland (2023).

<sup>7</sup>See, e.g., LaRock (2016: 151–80), LaRock & Collins (2016: 137–46), LaRock (2008: 201–22), Schwartz and Begley (2002), Guta (2011: 69–70), and Horgan (1993: 313–14).

<sup>8</sup>See, Perler (2015).

Moreover, a capacity gets its identity and proper metaphysical categorization from the type of property it actualizes.<sup>9</sup> The nature of a capacity-to-exemplify-F is properly characterized by F itself. Thus, the capacity to reflect light is properly considered a physical, optical capacity. For property dualists, the capacities for various mental states are mental and not physical capacities. Thus, the faculties that are constituted by those capacities are mental and not physical faculties. Multi-dispositional properties are classified by the range of properties they can actualize. Moreover, part of the nature of a substance resides in the dispositions and faculties that constitute it. This is one reason why SDists take debates about PD and the nature of these dispositions and faculties to have a direct bearing on the justification for SD.

A particular is the kind of thing it is in virtue of the actual and potential properties/faculties essential and intrinsic to it. Thus, a description of the faculties of a thing provides accurate information about the kind of particular that has those faculties. Moreover, a description of a particular's capacities/faculties is a more accurate source of information about its nature than is an analysis of the causal/functional conditions relevant for the particular to act in various ways. The latter can either be clues to the intrinsic nature of that particular or else information about some other entity that the particular relates to in exhibiting a particular causal action. Remember, there is a difference between attempts to describe, categorize, and identify a capacity's nature and possessor as opposed to proffering an explanation of the functional/causal conditions that must be present for that capacity to be actualized.

For example, if Smith needs to use a magnet to pick up certain unreachable iron filings, information about the precise nature of the magnet and its role in Smith's action does not tell us much about the nature of Smith (except that he is dependent in his functional abilities on other things, e.g., the magnet). We surely would not conclude that the actual and potential properties of a magnet are clues to Smith's inner nature. Similarly, functional dependence on/causal relations to the brain are of much less value in telling us what kind of thing a human person is than is a careful description of the kind-defining mental capacities/faculties, human persons as such possess.

Nevertheless, some physicalists may remain unmoved by my rejoinders. The following counterargument has been suggested to me: It is notable that either none or almost none of the philosophical phenomena for which I claim SD proffers a better explanation than physicalism would qualify as empirical discoveries. Maybe SD's ability to explain such things really does count in favor of the view, but if there are no empirically discovered psychological facts that SD is best able to predict or explain (e.g., chimerical colours, susceptibility to certain kinds of optical illusions, regular errors in reasoning or cognitive illusions, etc.), that would seem to show that science really does provide a reason (admittedly a defeasible reason) to prefer physicalist views over SD, and it would seem to go some way toward vindicating the view that there really is a problem for SD here: SD plays no role in predicting verified empirical data or explaining those data in our best scientific explanation of the findings of empirical psychology, while physicalist views do. Thus, this inadequacy provides a good reason to reject SD. Call this the explanatory/predictive problem. It is an aspect of the broader explanatory impotence claim.

What should we make of this problem? I do not think it carries enough weight to provide significant support to the explanatory impotence charge. To support my assessment, I will begin with two general responses to the problem, followed by three more specific defeaters that zero in on the explanatory/predictive inadequacy with respect to empirical data and empirical psychology.

First, philosophers of science have led the way in cataloging a dozen or so epistemic virtues which, when possessed by a theory, provide it with positive epistemic support, e.g., empirical adequacy or scope, epistemic simplicity, ability to solve internal and external conceptual problems, universal coherence (consistency with warranted beliefs acquired by direct experience).<sup>10</sup>

A problem surfaces when rival theories  $T_1$  and  $T_2$  are supported by different theoretical virtues. In such cases, a way to avoid treating  $T_1$  and  $T_2$  as incommensurable involves two things: (1) Recognize that the theoretical

<sup>9</sup>For multi-tracked dispositions, an actualized manifestation characterizes an essential constituent of the disposition's essence but is not exhaustive of it. We are using powers and dispositions interchangeably.

<sup>10</sup>Cf. Keas (2018).

virtues, the depiction of the nature and prioritization of the (alleged) facts to be explained, and the fittingness of those facts within the theory should be derived by the theory itself and not from its rival. (2) Reexamine the facts to be explained from each rival's perspective to see if one theory is (i) less ad hoc, (ii) not question-begging, (iii) better exhibits a fittingness or naturalness with the facts to be explained.

The notion of fittingness or naturalness is intuitive but, admittedly, not very precise. To tighten up things a bit, note that regarding naturalness, the types of entities postulated, along with the sorts of properties/powers they possess and the relations they enter should be “at home”—there should be a fittingness—with other entities in the theory. More precisely:

*Naturalness:* Some entity (a particular thing, process, categorical or dispositional property, or relation)  $e$  is natural for a theory  $T$ , just in case, either (i)  $e$  is a central, core entity of  $T$  or (ii)  $e$  bears a relevant similarity to central, core entities in  $e$ 's category within  $T$ .

For example, if  $e$  is in a category such as individual, force, property, event, relation, or cause, then  $e$  should bear a relevant similarity to central, core entities of  $T$  in that category. This is a formal definition, and the material content given to it will depend on the theory in question.

Moreover, given rivals  $R$  and  $S$ , the postulation of  $e$  in  $R$  is ad hoc or *question-begging* against advocates of  $S$  if  $e$  bears a relevant similarity to the appropriate entities in  $S$ , and in this sense, is “at home” in  $S$ , but fails to bear this relevant similarity to the appropriate entities in  $R$ .<sup>11</sup> The notion of “being *ad hoc*” is notoriously difficult to specify precisely. It is usually characterized as an intellectually inappropriate adjustment of a theory whose sole epistemic justification is to save the theory from falsification. Such an adjustment involves adding a new supposition to a theory not already implied by its other features. In the context of evaluating rivals,  $R$  and  $S$ , the principle just mentioned provides a sufficient condition for the postulation of  $e$  to be ad hoc and question-begging.

To tie things together, the nature of SD and PD entail that points (1) and (2) above (neither SD nor PD is primarily a theoretical postulate but a report of that of which we have direct awareness; as a theoretical postulate, the facts to be explained are distinctively philosophical) are the fundamental factors dualists employ to justify their views. As a result, empirical predictability/explanation of scientific data are not important epistemic virtues to SDists and PDists. Physicalists ought not to impose these theoretical virtues on their rival since dualism's very nature should determine its own ranking of theoretical virtues, the nature and prioritization of the data to be explained (the nature of phenomenal conscious states, their specific sort of wholistic unity, etc.), and so forth.<sup>12</sup>

Thus, SDists will not be troubled by its alleged inadequacy in predicting and explaining empirical data. Perhaps physicalists think their ability to more successfully predict and explain such data counts in favor of physicalism. But this would be a misunderstanding of the nature of evaluating SD/PD and physicalism as rivals. SD/PD should be evaluated by their own theoretical virtues and so forth. Thus, the value of predictive/explanatory success regarding empirical data counts for very little if SD/PD are to be evaluated—as they should—on their own terms and not in light of the preferred theoretical virtues of its rival.

A physicalist could respond that SD/PD has the wrong theoretical virtues, or at least the wrong prioritization of them. But this charge moves the debate into a second-order one about the relative value of different theoretical virtues for a theory (e.g., dualism). Moreover, the second-order debate as to which theoretical virtues and so on ought to be employed in the dualist-physicalist rivalry will itself be largely, if not strictly, a philosophical matter, e.g., should one adopt a first- or third-person perspective as the fundamental way of getting at the nature

<sup>11</sup>For example, suppose theory  $S$  explains phenomena in terms of discrete corpuscles and actions by contact, while  $R$  uses continuous waves to explain phenomena. If some phenomenon  $x$  was best explained in corpuscularian categories, it would be ad hoc and question-begging for advocates of  $R$  simply to adjust their entities to take on particle properties in the case of  $x$ . Such properties would not bear a relevant similarity to other entities in  $R$  and would be more natural and at home in  $S$ .

<sup>12</sup>For more on this, see Taylor (2019).

of phenomenal conscious and its possessor/unifier? If this is right, then assessing physicalist claims about the significant epistemic importance of physicalism's track record vis a vis dualism's in explaining or predicting empirical data is more complicated than merely citing cases where physicalism gains support from this issue. Assessing that importance must consider the nature of rival-theory evaluation along with weighing the epistemic significance of second-order issues.

There is a second general observation that adds strength to my first response. It concerns an important component in accurately assessing the strength and success of a theory, especially in light of anomalies or other problems that raise difficulties for that theory. Consider some theory T that could be present in a number of academic disciplines. Regarding T, there is a distinction between what I call *theory carriers* and *peripheral factors*. These differ primarily by the degree of epistemic ingression they exhibit for T. Theory carriers are fundamental, core factors of T such that if justified, they convey a great deal of epistemic support for T. By contrast, a peripheral factor is in some way implied by T, but its truth or falsity does not carry as much weight with respect to T as do the theory carriers.

If anomalies surface among the peripheral factors, their falsifying impact for T is not merely a factor of the anomalies considered in their own right. Rather, assessing their impact on T must take into consideration the strength of T's theory carriers. If the latter are strong, then the epistemic assessment of considering the peripheral factors as real versus alleged falsifiers must take into account all the evidence provided by the theory carriers. Why? Because if one takes these as real falsifying anomalies, one must abandon or adjust the theory, and that would be justified only if their combined weight was sufficient to overturn the epistemic support for T from its theory carriers. In many cases, it is more rational to take the anomalies as currently unexplained alleged anomalies.

Here is a concrete case from organic chemistry that will clarify what I've said. A standard organic chemistry text by Morrison and Boyd discusses a certain kind of chemical reaction known as halohydrogenation (Morrison and Boyd 1992). This kind of reaction was supposed always to happen in a certain way (relating to the order/location of placement of the hydrogen and the halogen in a hydrogen halide when it reacts with an alkene in accordance with Markovnikov's rule) based on scientific understanding of the theory carriers related to halohydrogenation.

However, a clear, refuting counterexample was found (hydrogen bromide) which, if judged simply by the evidence for that case in isolation from the evidence for a hypothesis's theory carriers, provided a refutation of the hypothesis and encouraged an abandonment of its theory carriers. But scientists suspended judgment on this case, treated the counterexample as an alleged anomaly, and engaged in ad hoc harmonization attempts from 1869 to 1933. At that time, they discovered new, hitherto unknown factors which allowed scientists to explain the peripheral anomaly without it having a negative impact on the theory and its theory carriers.

Why were they rational in doing this? Because the evidence for treating the counterexample as a real anomalous refutation had to be sufficiently strong to overturn the combined evidence for the theory carriers from a variety of sources that the hypothesis was true. If the counterexample was judged on its own, various interpretations that harmonized it with the hypothesis would have been ruled out. But this is not the proper course of action motivated by an accurate understanding of the epistemic situation. The counterexample was a member of a class of peripheral factors, and the evidence supporting the theory carriers was too strong to justify abandoning the theory in light of those peripheral factors.

Applied to assessing the epistemic impact on SD/PD of its empirical predictive/explanatory inadequacy vis a vis physicalist rivals, the dualist should argue that even if he/she grants that this inadequacy tends to count against dualism, it is a peripheral problem. As such, the negative impact of this inadequacy is far from sufficient to override the support dualism gets from its two theory carriers.

One could respond that if the number of peripheral anomalies increases, at some point, their combined weight justifies rejecting theory T and that is the situation regarding the large number of cases of relevant empirical predictive/explanatory data in support of physicalism over dualism.

I do think that such an objection works in the case of halohydrogenation theory regarding alkenes and hydrogen halides. If a whole range of different, relevant kinds of chemical reactions were anomalous, then at some

point, their combined weight would justify abandoning halohydrogenation theory. But I do not think it is successful when applied to dualism. Given the nature of (property or substance) dualism, the low epistemic value of empirical predictive/explanatory data is due to the *kind* of phenomena that characterizes these data. They are classified as peripheral due to the low epistemic value of this kind of problem regardless of how many cases of the kind are discovered.

So much for my general responses. I turn to three more specific defeaters that focus on dualism's explanatory/predictive inadequacy with respect to empirical data and empirical psychology. First, a number of such cases cited in favor of physicalism are actually empirically equivalent among physicalism (reductive or functional), mere property dualism, and substance dualism (SD). For example, let us assume that humans have mirror neurons such that if they are damaged, then one cannot feel empathy for another. How are we to understand this empirical phenomenon?

At least three empirically equivalent solutions are available: (1) strict physicalism (a feeling of empathy is identical to something physical, e.g., the firings of mirror neurons); (2) mere property dualism (a feeling of empathy is an irreducible state of consciousness *in the brain* whose obtaining depends on the firing of mirror neurons); (3) SD (a feeling of empathy is an irreducible state of consciousness *in the soul* whose obtaining depends while embodied on the firing of mirror neurons). Of these three, no empirical datum can help us decide which is correct. Simplicity and other non-empirical factors are beyond the scope of the current discussion.

In my opinion, the ubiquitous presence of such empirically equivalent cases helps to explain why three Nobel Prize winners working in neuroscience and related fields—John Eccles (substance dualist), Roger Sperry (mere property dualist) and Francis Crick (strict physicalist) could hold different ontologies regarding consciousness and the self even though they all knew the same neuroscientific data. While this fact could be explained by adopting, say, a Duhemian antirealist philosophy of science, that is not what is going on here. It is the fact of empirical equivalence that opens the door for the different ontologies, and this situation is consistent with a various sorts of scientific realism.

Consider a recent example offered by Andrew Y. Lee (2021, pp. 263–298). In a rigorous, innovative paper, Lee begins by stating and criticizing what he calls the Standard Framework for mental qualities with a specific focus on colours and colour experiences. According to the Standard Framework (or a quality–space model as a specific formulation of the Framework), the mental qualities of colour experiences are correlated with and represented by points in geometrical spaces such that distance between points inversely corresponds to degrees of phenomenological similarity.

Lee argues that the Framework is inadequate because it cannot capture *precision structure*, e.g., the contrast between seeing an object as crimson in foveal vision vs. seeing it as red in peripheral vision. So, Lee develops a new model to expand and replace the Standard Framework. Lee's model employs regions (sets of points) to model mental qualities. He also appeals to fields (a field on a space of points is an assignment of values to every point in the space) to accommodate perceptual confidence theory. A very important argument to which Lee appeals for his model is the fact that it generates new predictive hypotheses about the spaces and structures of mental properties constitutive of phenomenal colour experiences. And, he says, some of these predictions have been empirically verified and explained by his model.

Further details of Lee's position are not needed for the point I am making. Lee (2021, p. 4) explicitly states:

I will be neutral on most questions about the nature of mental qualities. In particular, I remain neutral on whether mental qualities can be instantiated even in the absence of consciousness, on the metaphysical relationship between mental qualities and perceptible qualities, and on whether mental qualities are fundamentally physical. Staying neutral on these issues ensures that the framework I develop is compatible with a wide variety of philosophical positions.

He concludes (Lee 2021, p. 35) by saying that



By formally modeling conscious experiences, we not only sharpen our understanding of how conscious experiences are structured, but also progress our understanding of consciousness without needing to resolve long-standing theoretical disputes about the mind-body problem, the nature of perception, or the physical correlates of consciousness.

In my view, this kind of project is one of the most promising ways of moving consciousness research towards a systematic science.

Lee's publication is a clear, recent, rigorous presentation of a model for empirical research regarding conscious experiences of colour that include important predictions of empirical data. Yet he explicitly does his work by setting aside philosophical issues regarding the nature of consciousness and the self. His model exhibits the empirical equivalence of physicalist, mere property dualist, and substance dualist interpretations of his model's empirical success. I think that successful explanations of psychological phenomena are generally as indifferent among physicalism, PD, and SD as is Lee's theory, even if proponents do not state or realize this.

My second specific defeater is the fact that SDists have, in fact, generated theories that have had predictive/explanatory success regarding empirical data. In notes 10–12 above, I have provided several sources for this claim. Here I offer a specific example.

In 2011, Jeffrey Schwartz and Rebeck Gladding (Schwartz and Gladding 2011) published a radical new theory for healing patients with obsessive compulsive disorder (OCD).<sup>13</sup> Prior to its publication, the standard approach to OCD was based on a physicalist (behaviorist) understanding of the human person and consciousness. Schwartz was dissatisfied with this approach due to its inadequate success and the frequent traumatization of patients subjected to it. So, he and Gladding proposed a new theoretical model that included three irreducible dualist elements: a substantial soul/self, libertarian agency and active power, and the existence of semantic, propositional content constitutive of thoughts.

They postulated that when an OCD patient faced an overwhelming urge to, say, wash his hands, since the soul transcended mere physical causation and laws of nature, the patient should pause, freely choose to entertain certain thoughts (i.e., semantic contents) as part of his self-talk, and habituate this activity for a few weeks. Brain scans were taken before and after this treatment was practiced, and their predictions were empirically verified. The regions of the brain associated with OCD were damaged prior to treatment but had been significantly healed after treatment. Thus, predictions based on the assumption that SD obtains were verified empirically, showing that SD is not explanatorily impotent with respect to empirical observations.

Admittedly, it is still true that there is not a large number of such cases. However, for two reasons I do not think this fact counts as a significant defeater for substance or property dualism. First, I have repeatedly noted that the vast number of dualists take (1) and (2) to be the theory carriers for dualism. Accordingly, their efforts have been directed in defending, clarifying, and responding to defeaters to them. As a result, little effort has been spent on the empirical aspects of dualism, though more work needs to be done here going forward.

Second, the sociology of practitioners in empirical psychology, neuroscience and related fields is such that employing dualism in theoretical and empirical work is eschewed by members of these communities. Nowhere is this more evident than in the methodological constraints on what counts as appropriate scientific investigation. As Jaegwon Kim noted, even though analytic and ontological behaviorism is no longer taken seriously, *Explanatory Methodological Behaviorism* is very much alive and well<sup>14</sup>:

1. *Explanatory Methodological Behaviorism*: Psychological theories must make no reference to inner *mental* states in formulating psychological explanations/predictions.

<sup>13</sup>For another example, see Schwartz et al. (2005).

<sup>14</sup>For the best brief treatment of Behaviorism, including different versions of methodological behaviorism, see Kim (2011).

With slight adjustments, this becomes *Explanatory Methodological Physicalism*:

2. *Explanatory Methodological Physicalism*: Theories in empirical psychology, neuroscience, and related fields must make no reference to inner *mental* states in formulating explanations/predictions of empirical data.

In (1), “inner” is not qualified and includes brain states that are spatially inside the brain and irreducible mental states that are non-spatially *in* the subject to which the subject has private access. In (2), “inner” allows for the former and disallows the latter.

Given the ubiquitous employment of this methodological constraint, it is hardly surprising that dualist theories are rare. Such theories are not considered scientific theories since they do not conform to proper methodology. Dualists must do more work on criticizing this constraint, and if they do, the issues will reside largely in philosophy of science. Still, the fact that this constraint is currently ensconced as *the* methodology for these fields, implies that dualist theories are regarded as non-scientific. It is not that dualist explanatory/predictive theories of empirical data have been weighed in the balance and found wanting. Rather, they have been disqualified from being weighed at all, at least for most practitioners.

My third specific defeater focuses on a different sort of case that allegedly provides straightforward evidence against PD and, thus, SD. The cases are not offered merely as examples where physicalist theories are superior to dualist rivals in explanatory/predictive success. Rather, the claim is made that their success strictly entails physicalism; only if physicalism is true will the theory have any such success in the first place. This is a strong position and the paradigm case of this sort of approach is one made by Paul Churchland (2005) regarding chimerical colours.

Churchland argues that if the Hurvich–Jameson opponent-process colour theory (Leo M. Hurvich and Dorothea Jameson 1955, pp. 602–616) is applied to chimerical colours, it can be used to show that the dualist view of subjective experience is entirely mistaken and false. He argues that the H–J theory can be used to predict and explain that there are chimerical colours that can be taken to be entirely physical properties. Hence, colour qualia are entirely devoid of subjective phenomenal properties. In light of this, Churchland argues that subjective experience should be type-reduced to neuronal activities.

As a rival to trichromatic theory, the H–J model is a theory according to which *opponent hues* cancel each other when superimposed: yellow and blue on the one hand, and green and red on the other. Hurvich and Jameson developed an experimental procedure called *hue cancellation*, which allowed them to quantitatively express the relative amounts of each of the four basic hues present in any spectral stimulus of a particular wavelength. Churchland employs the predictive/explanatory success of the H–J model, especially as it applies to chimerical colour experiences to argue for reductive physicalism as the reason the model succeeds. If Churchland is correct, then as noted earlier, his conclusion would render SD false.

For the following reasons, I remain unpersuaded by Churchland's arguments. First, there are good reasons to think that Churchland fails to type reduce phenomenal colour properties. In testing the H–J model, Hurvich and Jameson (Churchland 2005, pp. 544–553) acknowledge that their methodology required obtaining first-person, privately accessible reports about the phenomenal states subjects were in.<sup>15</sup> Now is not the time to rehearse why the fact that phenomenal states are fundamentally accessible by private, first-person

<sup>15</sup>Regularly, Churchland invites the reader to test H–J phenomenal colour experiences by performing certain exercises he provides. However, all of these irreducibly involve first-person private access to the qualia constitutive of the private experiences of the reader, experiences that must be reported to investigators since they have no access to them. Moreover, throughout Churchland's paper, he constantly claims—quite rightly—that the H–J model predicts and provides explanations for the physical factors that *cause* various private phenomenal colour experiences each with its own what-it-is-like. Churchland's closing section argues that the only conclusion we can justifiably reach is type identity. But his arguments for this (or against the possibility of inverted qualia) are all philosophical, e.g., arguments for the proper way to interpret intertheoretical reduction in the history of science, issues about a priori knowledge, and the philosophy of modality (especially metaphysical necessity vs. contingency).

introspective awareness supports property dualism. But it should be noted that their model is called a psychophysical one obtained by psychophysical methodology.<sup>16</sup> The chromatic response functions are determined experimentally for each individual observer, but, although there are slight interpersonal differences, the general shape of these functions remains the same when large numbers of subjects are considered. Chromatic response functions are assumed to correlate directly with hue *perception*. These slight interpersonal differences arise from the private accessibility of the phenomenal states being studied. Accordingly, some argue that there is no guarantee that the language one uses in talking about the imaginary/impossible colours is accurately describing what is being experienced.

Churchland explicitly acknowledges that chimerical colours are meant to be grounded in subjective visual experiences that people report to cognitive neuroscientists. And it is precisely the phenomenal subjective nature of these experiences that allegedly justify the claim that there are no qualitative colours in the external world to which these chimerical experiences correspond. The existence of experiences of impossible colour is being inferred based on the first-person reports of people who claim to have had them. That means that in this case, the first-person-based data are equally accessible to property dualist views of colour-conscious experience. So, I fail to see what special privilege the reductive physicalist views enjoy here. Indeed, arguably, the first-person nature of these reports seems to support PD and, as many have argued, SD.

Finally, let us grant that some people can experience impossible colours. Furthermore, let us assume that reductive physicalist views predict the reality of such experiences. But how can this give a unique advantage to reductive physicalist views with respect to predictive power? I ask this because property dualist views also can make such predictions without any problem. To do this, all they have to do is interview some people and see what they say about their experience of the colours in question. Based on the data they gather, defenders of the property dualist views can draw certain conclusions. Since this research is based on first-person data, I remain unconvinced that somehow, reductive physicalism has a unique way of successfully handling this matter.<sup>17</sup>

In sum, there are plausible rejoinders to the explanatory-impotence objection, and I have tried to present and clarify some of them.

## 2 | ACKNOWLEDGMENTS

I am deeply indebted to Mihretu Guta for comments on the issues surfaced by Churchland's article.

### ORCID

James Moreland  <https://orcid.org/0000-0002-4749-9692>

### REFERENCES

- Antonietti, A. (2008). Must Psychologists Be Dualists? In A. Antonietti, A. Corradini & E. J. Lowe (Eds.), *Psycho-Physical Dualism Today* (pp. 37–67). Lanham: Lexington Books.
- Bayne, T. (2018). Problems with Unity of Consciousness Arguments. In J. J. Loose, A. J. L. Menuge, & J. P. Moreland (Eds.), *The Blackwell Companion to Substance Dualism* (pp. 207–225). Oxford: Wiley-Blackwell.
- Bennett, M. R. & Hacker, P. M. S. (2003). *Philosophical Foundations of Neuroscience*. Malden: Blackwell Publication WWW color naming (n.d.). Retrieved October 10, 2022, from <http://cogsworth.cse.buffalo.edu/pub/colornaming/diss/subsection2.7.4.5.html>

<sup>16</sup>See [cogsworth.cse.buffalo.edu/pub/colornaming/diss/subsection2.7.4.5.html](http://cogsworth.cse.buffalo.edu/pub/colornaming/diss/subsection2.7.4.5.html).

<sup>17</sup>I am deeply indebted to Mihretu Guta for comments on the issues surfaced by Churchland's article. The same goes for an anonymous referee who painstakingly read the manuscript and gave me a detailed list of very helpful suggestions. This article is much improved due to his or her efforts. Finally, thanks go to my research assistant Eli Haitov for his suggestions on an earlier draft of this article and for formatting this for publication.

- Churchland, P.M. (2005). Chimerical Colors: Some Phenomenological Predictions from Cognitive Neuroscience. *Philosophical Psychology*, 18, 527–560.
- Churchland, P.M. (2013). *Matter and Consciousness*. Cambridge: MIT Press.
- Guta, M.P. (2011). Neuroscience or Neuroscientism? *Perspectives on Science and Christian Faith*, 63.1, 69–70.
- Hopp, W. (2011). *Perception and Knowledge*. Cambridge: Cambridge University Press.
- Horgan, T. (1993). Nonreductive Materialism and the Explanatory Autonomy of Psychology. In S. Wagner & R. Warner (Eds.), *Naturalism: A Critical Appraisal* (pp. 295–320). Notre Dame: University of Notre Dame Press.
- Hurvich, L. M., & Jameson, D. (1955). Some quantitative aspects of an opponent-colors theory. II. Brightness, saturation, and hue in normal and dichromatic vision. *Journal of the Optical Society of America*, 45(8), 602–616.
- Keas, M. (2018). Systematizing the Theoretical Virtues. *Synthese*, 195, 2761–2793.
- Kim, J. (2011). *Philosophy of Mind* 3rd edition. Boulder: Westview.
- LaRock, E. (2008). Is Consciousness Really a Brain Process? *International Philosophical Quarterly* 48.2, 201–222.
- LaRock, E. (2016). Neuroscience and the Hard Problem of Consciousness. In T. Crisp, S. Porter & G. T. Elshof (Eds.), *Neuroscience and the Soul: The Human Person in Philosophy, Science, and Theology* (pp. 151–180). Grand Rapids: Eerdmans.
- LaRock, E. & Collins, R. (2016). Saving Our Souls from Materialism. In T. Crisp, S. Porter & G. T. Elshof (Eds.), *Neuroscience and the Soul: The Human Person in Philosophy, Science, and Theology* (pp. 137–146). Grand Rapids: Eerdmans.
- Lee, A.Y. (2021). Modeling Mental Qualities. *Philosophical Review*, 130, 263–298.
- Lowe, E. J. (2008). *Personal Agency*. Oxford: Oxford University Press.
- Lycan, W. (2018). Redressing Substance Dualism. In J. J. Loose, A. J. L. Menuge, & J. P. Moreland (Eds.), *The Blackwell Companion to Substance Dualism* (pp. 22–40). Oxford: Wiley–Blackwell.
- Manzotti, T. & Moderato, P. (2014). Neuroscience: Dualism in Disguise. In H. Robinson & A. Lavazza (Eds.), *Contemporary Dualism* (pp. 81–97). New York: Routledge.
- Morrison, R.T. & Boyd, R. N (1992). *Organic Chemistry* 6th edition. Englewood Cliffs: Prentice Hall.
- Murphy, N. (2006). *Bodies and Souls, or Spiritual Bodies?* Cambridge: Cambridge University Press.
- Murphy, N. (1998). “Human Nature: Historical, Scientific and Religious Issues.” In W. S. Brown, N. Murphy & H. N. Malony (Eds.), *Whatever Happened to the Soul?* (pp. 1–29). Minneapolis, Minnesota: Fortress Press.
- Nida-Rumelin, M. (2010). An Argument from Transtemporal Identity for Subject Body Dualism. In G. Bealer & R. Koons (Eds.), *The Waning of Materialism* (pp. 191–211). Oxford: Oxford University Press.
- Perler, D. (2015). *The Faculties: A History*. Oxford: Oxford University Press.
- Ravenscroft, I. (2018). Why Reject Substance Dualism. In J. J. Loose, A. J. L. Menuge, & J. P. Moreland (Eds.), *The Blackwell Companion to Substance Dualism* (pp. 267–282). Oxford: Wiley Blackwell.
- Rickabaugh, B., & Moreland, J. P. (2023). *The Substance of Consciousness*. Oxford: Wiley Blackwell.
- Rockwell, W. T. (2005). *Neither Ghost Nor Brain*. Cambridge: MIT Press.
- Schwartz, J.M., & Begley, S. (2002). *The Mind and the Brain: Neuroplasticity and the Power of Mental Force*. New York: Harper Collins.
- Schwartz, J.M., & Gladding, R. (2011). *You are Not Your Brain*. New York: Penguin Group.
- Schwartz, J.M., Stapp, H. P. & Beauregard, M. (2005). Quantum Physics in Neuroscience and Psychology: A Neurophysical Model of Mind–Brain Interaction. *Philosophical Transactions of the Royal Society B* 360, 1309–1327.
- Taylor, D. S. (2019). Two Kinds of Explanatory Integration in Cognitive Science. *Synthese*, 198, 4573–4601.
- Uttal, W. R. (2001). *The New Phrenology: The Limits of Localizing Cognitive Processes in the Brain*. Cambridge: MIT Press.
- Uttal, W. R. (2004). *Dualism: The Original Sin of Cognitivism*. Mahwah: Lawrence Erlbaum Associates.

**How to cite this article:** Moreland, J. (2023). On the alleged explanatory impotence/conceptual vacuity of substance dualism. *Ratio*, 36, 180–191. <https://doi.org/10.1111/rati.12372>