

## Naturalistic Dualism and the Problem of the Physical Correlate

Jonathan Schaffer, Rutgers University  
 Draft of June 21, 2020

It is sometimes asked, “Why can't there be psycho-physical laws which are of a novel sort, just as the laws of electricity and magnetism were novelties from the standpoint of Newtonian mechanics?” Certainly we are pretty sure in the future to come across new ultimate laws of a novel type, but I expect them to relate simple constituents: for example, whatever ultimate particles are then in vogue. I cannot believe that ultimate laws of nature could relate simple constituents to configurations consisting of perhaps billions of neurons (and goodness knows how many billion billions of ultimate particles) all put together for all the world as though their main purpose in life was to be a negative feedback mechanism of a complicated sort. Such ultimate laws would be like nothing so far known in science. (Smart 1959: 143)

The *naturalistic dualist* (Chalmers 1996, Nida-Rümelin 2007, Pautz 2009, Gertler *forthcoming*) proposes to explain the empirical correlations between physical states and mental states by positing *fundamental laws of psychophysics*. She can thus embrace the dualistic conclusions of the explanatory gap arguments while offering nomological explanations for the correlations, so promising a scientifically acceptable dualism.

I argue—building on Smart (1959), Seager (1995), and Latham (2000)—that the naturalistic dualist is positing fundamental laws with an *insufficiently natural physical correlate*. For an initial illustration, consider a posited fundamental law linking C-fiber firing to pain. The problem at issue is that C-fiber firing is a higher level property of middling scale organisms, and so unfit for a fundamental law. (The problem is not that these purported fundamental laws invoke mental states on the right, but rather with the sorts of physical states they must invoke on the left.)

More generally, a fundamental law of psychophysics must have a physical side that is both *sufficiently natural* to fit a fundamental law, and *humanly correlated* to explain the correlations. The problem is that the physical correlates of human mental states are not particle states or anything sufficiently natural on the physical side. One can visualize this as an *abutment problem* in bridge construction. A fundamental bridge law should abut unified ground level conditions:



But a neural (or functional, or informational) state is *too high* for the box, and its physical realizers are *too widely disjunctive* for a box:



There is no unified ground level property to box on the left.

If the physical correlate is insufficiently natural for a fundamental law, then the naturalistic dualist's proposed explanation of the correlations fails, and generally it seems that no form of dualism can explain the correlations. So I conclude that dualism looks like an explanatory failure. (I am not saying that the dualist is

failing to explain why there are mental states, but rather that she is failing to explain why our mental states twist in the neural flux.)

The problem of the physical correlate also serves—in a wider dialectical context—to favor the *ground physicalist* alternative (Schaffer 2017a, *forthcoming*), on which the correlations are understood via vertical ladders of grounding instead of horizontal bridges of causation. Though there are issues arising for ground physicalism too, and in any case one need not accept ground physicalism as the solution (or even admit a notion of grounding), to see that there is a problem for naturalistic dualism.

*Overview:* In §1 I describe naturalistic dualism and its fundamental laws. In §2 I develop the problem of the physical correlate. I conclude in §3 with reflections on the physicalism-dualism debate.

## 1. Naturalistic Dualism

### 1.1 Naturalistic dualism, tame and wild

The *naturalistic dualist* proposes to explain the empirical correlations between physical states and mental states by positing *fundamental laws of psychophysics* (FLOPs), so promising a scientifically acceptable dualism. In this vein Chalmers (1996: 127; see also 2003: 124) says:

Where we have new fundamental properties, we also have new fundamental laws. Here the fundamental laws will be *psychophysical laws*, specifying how phenomenal (or protophenomenal) properties depend on physical properties... [T]hey will be *supervenience laws*, telling us how experience arises from physical processes.

Likewise Nida-Rümelin (200: 270) defends a “dualist emergentism” which posits “psychophysical laws that are... fundamental laws of nature,” Pautz (2009: 64–6) defends a primitive consciousness relation which he argues favors a property dualism on which consciousness “supervenes on the physical with only nomological necessity,”<sup>1</sup> and Gertler (*forthcoming*: §1.3) writes:

[N]aturalistic dualists generally allow that, because of contingent laws of nature linking consciousness to structure and dynamics, consciousness will not actually vary independently of structural-dynamic phenomena. Dualism is compatible with the idea that conscious experience arises from structural-dynamic phenomena in a lawlike way.

Backing up a moment, I take *dualism* to be the view that the physical and the mental are both fundamental. I focus on *property dualism* (Chalmers 1996: 123–29), as motivated by zombies and other *explanatory gap arguments* (Chalmers 1996: 93–122).<sup>2</sup> And I assume that there are robust empirical correlations between physical and mental properties, involving what Crick & Koch (1998) call “the neural correlates of consciousness.” In this setting it is natural to think of the dualist-physicalist divided in terms of whether these correlations are merely nomological connections between distinct properties subserved by FLOPs, or more robust metaphysical connections subserved by grounding principles or reductive identities. In this vein Kriegel (*forthcoming*) summarizes the divide as:

[Dualism] is the view (roughly) that the experiential and the physical are mutually (metaphysically) independent, such that any links between consciousness and its neural correlate are at most causal and contingent, not constitutive and necessary.

---

<sup>1</sup> Though Pautz (2009: 66) is also open to a physicalism with primitive necessary connections—which he later (2017: 389) labels “grounding without reduction”—and doubts that anything hangs on the physicalism-dualism debate at this stage, claiming that the main issue is reductionism not physicalism.

<sup>2</sup> The main alternative is of course a Cartesian *substance dualism* on which the mental and physical are themselves understood as different individuals (e.g. a body and a spirit). It is possible but complicating to have a substance dualist version of naturalistic dualism (see fn. 3).

And Bennett (*forthcoming*) acknowledges that such a dualist “maintains a reasonable respect for the physical sciences, while simultaneously claiming that phenomenal properties are genuinely new additions to the world.” So naturalistic dualism can seem like a motivated and plausible scientific approach, which does not ignore the existence of neural correlates of consciousness, but explains the correlations nomologically.

It is worth distinguishing between *tame* and *wild* naturalistic dualisms. The tame version aims to be fully conservative and neutral on auxiliary matters of physics, mentality, and lawhood, and claims that we can explain the correlations in one clean stroke *just* via *FLOPs* (no fiddling with anything else). Tame naturalistic dualism is suggested by passages in Chalmers such as (1995: 20 [*italics added*]; see also 1996: 126 and 245):

The new basic principles postulated by a nonreductive theory give us the extra ingredient that we need to build an explanatory bridge... Nothing in this approach contradicts anything in the physical theory; *we simply need to add* further bridging principles to explain how consciousness arises from physical processes.

The wild version—in contrast—does not just posit *FLOPs* but also packages them with controversial speculations on auxiliary matters such as physics, consciousness, and lawhood. Chalmers is overall best read as being open to both tame and wild options, as he is open (1996: 287) to packaging naturalistic dualism with conceptions of physics on which “information itself is fundamental,” views of quantum mechanics (in Chalmers & McQueen *forthcoming*) on which consciousness collapses the wave function, and micropsychist views (1996: 299) on which the fundamental (proto-)phenomenal properties inhere in particles.

Tame naturalistic dualism is of course the most plausible naturalistic dualism. To the extent that the naturalistic dualist packages her *FLOPs* with still further controversial speculations, the plausibility of her resulting view is thereby compromised. (And in some cases her claim of scientific acceptability may be forfeit, and her claim to avoid an explanatory gap may be lost.) I do not mean to dismiss speculative views out of hand, especially in dealing with such a recalcitrant problem as the mind-body problem. A wild naturalistic dualism may be true in the end. I am only saying that it must be regarded as not so plausible for now.

I flag the distinction between tame and wild naturalistic dualisms because the naturalistic dualist has some escapes from the problem of the physical correlate (§2), but all require forays into the wild.

### 1.2 Fundamental laws of psychophysics

The naturalistic dualist treats physical and mental properties as both fundamental, positing *FLOPs* to explain the correlations. We can be more precise about *FLOPs*, schematizing them via a fundamental law operator over a universally quantified conditional linking physical states to mental states, as per:

*FLOP schema*: [F-law]  $(\forall x) (\textit{Physical correlate } x \rightarrow \textit{Mental property } x)$

This is to be read as “It is a fundamental law that anything in \_\_\_\_\_ is in \_\_\_\_\_,” where the second blank is filled by a mental property and the first by its physical correlate.<sup>3</sup> For instance, supposing just for illustration that C-fiber firing is the neural correlate of pain, the naturalistic dualist may posit laws such as:

*FLOP neural-pain*: [F-law]  $(\forall x) (\text{C-fiber firing } x \rightarrow \text{Pain } x)$

---

<sup>3</sup> *FLOP schema* is a property dualist schema, with physical and mental properties attributed to a common individual. For a substance dualist variant, we first need to introduce a *tethering* relation pairing bodies and spirits, and then formulate *FLOPs* as linking physical states to mental states across tethered body-spirit pairs:

*FLOP schema, substance*: [F-law]  $(\forall x)(\forall y) (\textit{Physical correlate } x \ \& \ \textit{Tethered } (x, y) \rightarrow \textit{Mental property } y)$

The property dualist needs no such tethering relation, since she can get by with mere co-instantiation in a common individual as what “tethers” physical and mental properties.

In English: “It is a fundamental law that anything that has C-fiber firing has pain.”<sup>4</sup>

I am not saying that *FLOP schema* is the only way, or even the best way, to precisify what the naturalistic dualist posits. I am just saying that it useful, both for clarifying naturalistic dualism and for pinpointing where objections arise. (The reader who thinks that the problem of the physical correlate would dissolve with a better formalism is invited to show how.)

My focus is on the “*Physical correlate*” slot in the antecedent position in *FLOP schema*, so I will largely take the remaining ideology for granted. But briefly: the fundamental law operator “[F-Law]” may be understood in terms of virtually any of the many views about laws of nature, for instance via *sparse second-order necessitation universals* (Armstrong 1983), or *axioms of the best system* (Lewis 1983), or *primitive nomic ingredients of reality* (Maudlin 2007), etc. I think that there is a scientific notion of a fundamental law, captured in Mill’s (1843: 41) talk of “ultimate laws” versus “derivative uniformities,” and exemplified by historical candidates such as Newton’s  $F=ma$  and Schrödinger’s equation. The “[F-Law]” operator is intended to capture this scientific notion in a metaphysically neutral way. Or at least, since the naturalistic dualist posits certain fundamental laws, I beg no questions by gifting her the ideology of “It is a fundamental law that...”

With respect to the “*Mental state*” slot in the consequent position, I use pain as a working example of what might go in here. But even within pain states, there is a question of whether the law should invoke human pain specifically, or a multiply realizable state of pain which even a mollusk, Martian, or machine might experience, or perhaps even some sort of (proto-)pain only experienced by individual particles. I will consider such options in due course.

Just to reiterate, I am not objecting to the fact that *FLOPs* invoke mental properties (e.g., pain) on the right. It would beg the question against the dualist to object to *that*. Rather I am objecting to the sorts of physical properties (e.g., C-fiber firing) that must be invoked on the left.

## 2. The Problem of the Physical Correlate

So far I have clarified naturalistic dualism, as positing that physical and mental properties are both fundamental but linked by *FLOPs*, usefully schematized via:

*FLOP schema*: [Fund-law]  $(\forall x) (\textit{Physical correlate } x \rightarrow \textit{Mental state } x)$

My question for the naturalistic dualist is, what goes in the “*Physical correlate*” slot on the left?

The problem of the physical correlate may be understood as the problem of simultaneously satisfying two guiding constraints:

*Sufficiently natural*: The physical correlate cannot be a higher level or widely disjunctive property

*Humanly correlated*: The physical correlate must be properly correlated with human mental states

The rationale for *Sufficiently natural* is that it represents an independently plausible constraint on fundamental laws, and the rationale for *Humanly correlated* is to enable an explanation for the correlations involving our

---

<sup>4</sup> I use “C-fiber firing” just for familiarity. Actually C-fibers (and A $\delta$ -fibers) are nociceptors in the peripheral nervous system, which signal potential trauma to the spinal cord. There can be pain without nociception, such as with “phantom limb” pain. A more modern perspective, initiated by the GATE theory of pain (Melzack & Wall 1965) and the neuromatrix model (Melzack 1999), moves away from seeing pain as rooted in a specific local activation of the peripheral nervous system, and towards seeing pain as rooted in a more holistic pattern of activation through a connected “Pain Matrix” in the central nervous system, involving areas such as the primary and secondary somatosensory cortices (S1 and S2), the insula, and the anterior cingulate cortex (ACC). For recent discussion see Garcia-Larrea & Bastuji 2018. I trust that the informed reader can make the empirically apt substitutions.

mental states. The problem is that these constraints clash. What is humanly correlated does not look sufficiently natural for a fundamental law. The “neural correlates of consciousness” are of course *neural*, and so too high level for a fundamental law.

We can be more precise about the problem, via the following argument. Let “*PC*” be any proposed entry into the “*Physical correlate*” slot for any proposed *FLOP*:

1. *PC* must satisfy *Sufficiently natural*
2. *PC* must satisfy *Humanly correlated*
3. Nothing satisfies both *Sufficiently natural* and *Humanly correlated*
4. If nothing satisfies both *Sufficiently natural* and *Humanly correlated*, then nothing can serve as *PC*
5. If nothing can serve as *PC*, then there are no *FLOPs*
6. Thus there are no *FLOPs*

The argument is evidently valid, and 4-6 should pass without comment. So it remains to discuss 1-3, as “where the action is” in the problem of the physical correlate.

### 2.1 Premise 1, the need to satisfy *Sufficiently natural*

Premise 1 of the problem of the physical correlate says:

1. *PC* must satisfy *Sufficiently natural*

Where (again) *PC* is any proposed entry into the “*Physical correlate*” slot for any proposed *FLOP*, and *Sufficiently natural* is the constraint against higher level or widely disjunctive properties appearing in fundamental laws:

*Sufficiently natural*: The physical correlate cannot be a higher level or widely disjunctive property

The core rationale for 1 is that it is an independently plausible constraint on fundamental laws that they involve no higher level or widely disjunctive properties, since such properties lack sufficient unity to engage with the fundamental nomological machinery. This encodes the idea that a fundamental bridge law should abut unified ground level conditions.

In more detail, the ban on invoking higher level properties in *PC* is written into most metaphysical accounts of lawhood, which often encode the even stronger condition that fundamental laws only link fundamental properties. To borrow terminology from Hicks & Schaffer (2017: 412), this is:

*Link*: Only fundamental properties can be invoked in fundamental laws

In this vein, Armstrong (1983) posits sparse first-order universals, and treats fundamental laws via second-order necessitation universals over the sparse first-order universals. Lewis (1983: 368) posits perfectly natural properties, and treats fundamental laws as axioms of the best systematization in a perfectly natural language: “Fundamental laws, those that the ideal system takes as axiomatic, must concern perfectly natural properties.” And North (2013: 186) says about the fundamental dynamical laws that they “relate what’s fundamental to what’s fundamental,” which is “why they are a guide to the fundamental nature of the world.” Relatedly Chalmers (2003: 124) says that there are fundamental laws “where there are fundamental properties.”

But there are more liberal options. For instance, Loewer (2007) allows the practice of physicists to play a role in determining eligibility. And Hicks & Schaffer (2017) argue that the practice of physicists is to permit some trading off of naturalness with other virtues, allowing precisely defined derivative properties to

feature in fundamental laws given sufficient gains to the resulting system of equations.<sup>5</sup> Even on these more liberal options, it should be evident that it is no part of the practice of physicists to invoke higher level properties. So one only needs the more modest principle:

*Weak link:* Higher level properties cannot be invoked in fundamental laws

There is a basis in scientific practice for *Weak link*, visible in the rejection of von Neumann's (1955) interpretation of quantum mechanics, on which 'measurement' is used to distinguish when the wave function evolves by Schrödinger's equation, from when it collapses by Born's rule. It is now widely acknowledged—as emphasized by Bell (1990)—that 'measurement' has no place in fundamental laws.<sup>6</sup> An even better analogy would be with imagined interpretations of quantum mechanics on which wave function collapse is triggered by “the presence of life,” or—to toss in some neural terminology—“the activation of the primary visual cortex,” or—to toss in some functional terminology—“the activation of nociceptors.” Such imagined interpretations deserve to be dismissed, since—given what we know about “the physical side” of the world (and leaving open whether it has a separate “mental side”)—higher level realized properties like measurement, life, visual cortex activation, and nociception have no place in fundamental laws.

But it is not enough merely to ban higher level properties from fundamental laws, since there is always the option of replacing the higher level property with a disjunctive property, disjoining all of its possible physical realizers. Let us imagine—just for definiteness—that the fundamental physical level involves *particles*, and in particular that fundamentally there are just particles with intrinsic mass and charge occupying points in spacetime. The key thing to note is that there will almost certainly be a vast plurality of diverse particle arrangements that can realize a higher level property like the neural property of C-fiber firing. (Just think of all the different persons—and stages of persons—that have experienced pain. Surely there were not all in precisely the same particle arrangement every time.) In general, the fundamental physical image of a higher level realized property is a widely disjunctive property, which in the case of *FLOP neural-pain* would be:

*FLOP realizers-pain:* [Fund-law]  $(\forall x) ((\text{Fund-phys1} \vee \text{Fund-phys2} \vee \text{Fund-phys3} \vee \dots) x \rightarrow \text{Pain } x)$

Here each 'Fund-phys $\underline{n}$ ' term refers to the  $\underline{n}$ th fundamental physical realizer of C-fiber firing, or whatever the physical correlate of pain might be, from some list of particle arrangements, or whatever the fundamental physical realizers might be.<sup>7</sup>

---

<sup>5</sup> For instance, Hicks & Schaffer (2017) focus on the idea that resultant force and acceleration are derivative properties in Newtonian mechanics (resultant force is the sum of the component forces, acceleration is the second derivative of position), even though Newton's  $F=ma$  is *the* historical paradigm of a fundamental law. They argue that the resulting equations are sufficiently elegant and modular to be worth some derivativeness.

<sup>6</sup> Relatedly, Gottfried (quoted in Bell 1990: 37) speaks of von Neumann's “infamous postulate” on which “the measurement act ‘collapses’ the state,” calling it “an ugly scar on what would be a beautiful theory...” Note that there are actually multiple reasons to reject ‘measurement’, including that it is higher level and also that it is imprecise. I am focused on the former reason, but accept the latter as well. These reasons should be separated. It may be that certain chemical level properties turn out to have unique and precise physical level realizations, but even then such chemical properties would have no place in fundamental laws.

<sup>7</sup> Note that I am not here taking a stand on multiple realizability across species. The disjunctive property in *FLOP realizers-pain* is for *human C-fiber firings* (or more generally, for the *human* correlate of pain), and is neutral on whether mental states like pain are species-specific (see Kim 1998), or realizable by mollusks, Martians, and machines (Putnam 1975). To include mollusk, Martian, and machine pains here—as I think one should—would merely engender an even wider disjunctivity.

In proposals like *FLOP realizers-pain*, the physical correlate is written with only perfectly natural predicates. But the physical correlate still fails to be sufficiently natural, not for invoking the wrong predicates, but for disjoining its (right) predicates. This ban on disjunctive properties may be understood via:

*No disjunctions*: Disjunctive properties cannot be invoked in fundamental laws

For instance, Armstrong (1978) would say that disjunctions of universals are not themselves universals, so that disjunctive conditions cannot stand in either the ‘*F*’ or ‘*G*’ slots of his second-order necessitation laws of the form Nec (*F*, *G*).

One may be more liberal than *No disjunctions*. For instance, laws with disjunctive properties may “enter the contest” for Lewisian best systemhood, but would presumably pay a penalty in simplicity. Kim (1998: 318) says that the problem is that widely disjunctive properties are not *projectible*, insofar as there is no reason to expect, from observing the effects of a given realizer Fund-phys<sub>*i*</sub>, that a fundamentally distinct realizer Fund-phys<sub>*j*</sub> (where *i* ≠ *j*) would have similar effects. A different approach—rooted in Putnam’s (1975: 295–98) comments on explaining why a square peg cannot fit through a round hole—would be to say that widely disjunctive properties are not *explanatory*, for failing to unify the phenomena. But in any case one can weaken *No disjunctions* to:

*No wide disjunctions*: Widely disjunctive properties cannot be invoked in fundamental laws

Where ‘width’ is left intuitive, encoding a to-be-specified measure involving the number and dissimilarity of the disjuncts, on the understanding that *FLOP realizers-pain* should count as widely disjunctive on any reasonable measure, given the number and dissimilarity of the particle arrangements that might subserve a human C-fiber firing, or more generally provide the physical image of any higher level neural property.

Premise 1 thus encodes the widespread and plausible claims of *Weak link* and *No wide disjunctions*, thereby ruling out options like *FLOP neural-pain* and *FLOP realizers-pain*, for featuring insufficiently natural properties on the physical side. It is of course open to the naturalistic dualist to simply reject *Weak link* and/or *No wide disjunctions*. But this involves fiddling with standard conceptions of lawhood, and so requires a first foray into the wild. Moreover, the resulting view strikes me as compromising the naturalistic dualist’s claim to a scientifically acceptable dualism, not for invoking mental properties in fundamental laws, but rather for rejecting the scientific image of fundamental laws as abutting unified ground level conditions.

## 2.2 Premise 2, the need to satisfy *Humanly correlated*

Premise 2 of the problem of the physical correlate says:

2. *PC* must satisfy *Humanly correlated*

Where (again) *PC* is any proposed entry into the “*Physical correlate*” slot for any proposed *FLOP*, and *Humanly correlated* is the constraint that the correlate must cover the human case:

*Humanly correlated*: The physical correlate must be properly correlated with human mental states

The core rationale for 2 is that the empirical correlations that need to be explained are between the physical and the mental states of middling scale organisms like us. As Goff (2009: 290) notes:

It is this kind of conscious experience, the conscious experience that corresponds to organisms, the kind of conscious experience that in our own case we are immediately acquainted with, that we want a theory of consciousness to explain. This is because this is the only kind of conscious experience that we have pre-theoretical reason to believe in.

A proposed *FLOP* that violated *Humanly correlated* could not suffice to explain the known correlations.

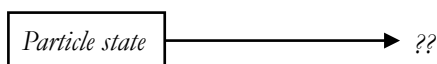
Note that *Humanly correlated* allows that the human cases may be covered in *neural* terms (e.g. C-fiber firing, as in *FLOP neural-pain*), more abstractly in *functional terms* (e.g. nociception), or also more abstractly in *informational terms* (e.g. having a certain informational shape). *Humanly correlated* also allows that the human case may be covered in terms of the reductive physical image of any of these proposals, such as in vast disjunctions of conjunctions of particles arrangements, each conjunct of which is a realizer of a given neural state (as in *FLOP realizers-pain*), or functional state, or informational state. There is a further question as to the best terms in which to theorize about the human case, on which I remain neutral.<sup>8</sup>

By way of exhibiting a physical correlate violating *Humanly correlated*, consider a proposed *micropsychist FLOP* which only correlates the physical and the (proto-)phenomenal states of individual *particles*, such as:

*FLOP particle-pain*: [Fund-law]  $(\forall x)$  (Unit negative charge  $x \rightarrow$  Pain  $x$ )<sup>9</sup>

This law has unit negative charge in the “*Physical correlate*” slot, so mapping electron charges to electron pains. *FLOP particle-pain* thus falls silent on human pains, and so cannot explain why human pain correlates with anything, much less anything like C-fiber firing or nociception.

Indeed micropsychism is especially relevant given the combination of *Sufficiently natural* (§2.1) and our working supposition that the fundamental physical level concerns particles.<sup>10</sup> For—in this setting—the physical correlates seen in micropsychist proposals like *FLOP particle-pain* are precisely what would allow *FLOPs* to abut unified ground level conditions. In other words, if we start from the idea that the unified ground level physical states are individual particle states, then *Sufficiently natural* demands that *FLOPs* must have individual particle states as their antecedents:



Now it is very hard to avoid concluding that what fills in for the question marks must be individual mental states of particles. Indeed this follows immediately from the sort of property dualism built into *FLOP schema*, since if the law embeds a universally quantified conditional of the form ‘ $(\forall x)(\Phi x \rightarrow \Psi x)$ ,’ where ‘ $\Phi x$ ’ is a particle physical state and so ‘ $x$ ’ a particle, then ‘ $\Psi x$ ’ can at most be a mental state of *that particle*.<sup>11</sup>

One can thus see micropsychism as a kind of *reverse-engineered solution* to the problem of satisfying *Sufficiently natural* (given a particle-based view of physics). Indeed something like this view for something like this reason is tentatively explored by Chalmers (1996: 293–99) and reluctantly endorsed by Seager (1995: 280), on grounds that if consciousness neither reduces to anything physical, nor emerges only from complex configurations, then “elements of consciousness must be found in the basic construction materials of the

<sup>8</sup> I think that the correlations are best understood in functionalist terms (Schaffer *forthcoming*; see also Chalmers 1996: ch. 7). But the problem of the physical correlate arises independently of whatever view one takes on the matter.

<sup>9</sup> Note that the intended notion of “charge” is the fundamental notion of *gross charge*, as opposed to the derivative notion of *net charge derived from summing the gross charges*. Electrons have both, but composite human organisms have only net charge, as a result of being composites of particles with various gross charges.

<sup>10</sup> On a holistic view of the fundamental physical level (e.g., one invoking the quantum wave function of the cosmos), one should instead consider *cosmopsychism* (see Shani 2015, Nagasawa & Wäger 2016, Goff 2017, and Albahari 2020), with *FLOPs* mapping physical states to (proto-)phenomenal states of the cosmos.

<sup>11</sup> On our substance dualist variant (fn. 3), ‘ $\Psi$ ’ can at most be a mental state of the spirit tethered to that physical particle. Hardly progress!



universe.” Seager (1995: 286) thus concludes that micropsychism “is the most natural way to incorporate consciousness as truly fundamental.”

The point is not that micropsychist proposals such as *FLOP particle-pain* go wrong in positing mental states for particles (I suspect that they do go wrong in this way, but that is not the point at issue). Rather the point is that such micropsychist proposals go wrong in not connecting to the mental states of humans.

That said, there is a principled way in which the naturalistic dualist could reject *Humanly correlated*, namely by positing still further principles—in addition to her *FLOPs*—serving to correlate the outputs of her *FLOPs* onwards to human mental states. For instance, she might embrace micropsychist *FLOPs* such as *FLOP particle-pain*, while tacking on further *combination principles* connecting the mental states of particles to the mental states of the wholes they compose, so claiming a “two-step” explanation for the known correlations.

Of course such a view requires two forays into the wild, both in attributing consciousness to things other than middling scale organisms like us (such as particles), and in tacking on further principles (such as combination principles) to make an explanatory connection to human mental states.

But even that is not enough, for a dilemma still lurks with respect to these further principles (e.g. combination principles), as to whether they are to be understood as further fundamental laws of nature, or something stronger such as grounding principles or reductive identities. If the further principles are held to be further fundamental laws of nature, then these merely seem to (re-)generate comparably bad problems to the problem of the physical correlate. For now one needs to posit something like Fundamental Laws of Combination (*FLOCs*) linking, e.g., individual particle (proto-)mental states to the mental states of middling scale organisms like us. And the fundamental (proto-)mental image of a human mental state will still be a complex and multiply realizable matter, meaning we will either need a unified but higher level description or a vastly disjunctive lower level description of it. This seems to me just to move the bump under the rug, replacing a problem with *FLOPs* with a comparable problem with *FLOCs*.

On the other hand, if the combination principles are held to be something stronger than fundamental laws, such as grounding principles or reductive identities, then allowing such principles seems to undermine the very explanatory gap arguments that led away from physicalism. For first, as Seager (1995: 280–81), Goff (2009), and Chalmers (2016) have discussed, we arguably face the very same sort of explanatory gap worries getting from the micro-phenomenal to the macro-phenomenal as we do in getting from the physical to the macro-phenomenal. As Goff (2009) puts the point, just as it is conceivable that there are *zombies* who are physical duplicates of us but lack conscious experience, so too is it conceivable that there are *microexperiential zombies* which are physical and particle-experience duplicates of us but lack conscious experience. Just as the physical state of an organism fails to *a priori* entail any particular state of consciousness for that organism, so the physical state of an organism plus the microexperiences of its particulate constituents fails to *a priori* entail any particular state of consciousness for that organism. As Stoljar (2006: 120) says: “[I]t seems just as hard to see how one experiential truth can entail another as it is to see how a nonexperiential truth can entail an experiential truth.” If the naturalistic dualist is allowing herself to posit grounding principles or reductive identities to bridge the explanatory gap from the micro-phenomenal to the macro-phenomenal, how can she forbid the physicalist from positing grounding principles or reductive identities to bridge the explanatory gap from the physical to the macro-phenomenal?<sup>12</sup>

### 2.3 Premise 3, the clash

Premise 3 of the problem of the physical correlate says:

3. Nothing satisfies both *Sufficiently natural* and *Humanly correlated*

---

<sup>12</sup> For an argument that grounding principles are needed anyway, and generally bridge explanatory gaps, see Schaffer 2017a. For a reply see Rabin 2019.

Here is where the problem of the physical correlate comes to a head. The co-constraints on the physical correlate (§§2.1–2.2) are not co-satisfiable but clash. The core of the issue is that, as a matter of empirical fact, mental states are known to be found in middling scale organisms like us, whose physical image is not sufficiently natural for fundamental laws.

One can approach the clash from either side. Starting from *Sufficiently natural*, what is wanted—as exemplified by the pressures towards micropsychism discussed in §2.2—is to take the antecedents of *FLOPs* to concern, not insufficiently natural higher level or widely disjunctive properties, but unified ground level states. Depending on how one thinks of the fundamental physical side, these antecedents might concern the states of individual particles or the states of the whole cosmos, but then the consequents of such *FLOPs* will concern the (proto-)mental states of individual particles or of the whole cosmos. In neither case will we satisfy *Humanly correlated*.

Generalizing the discussion in §2.2, we can think of it this way:



Now the only further claim needed is the vastly plausible empirical claim of:

*Human disunity*: The physical correlates of human mental states are not sufficiently natural

This encodes the idea that we are middling scale organisms, and that “the physical side” of the world (leaving open whether it has a separate “mental side”) offers no fundamental properties concerning states of middling scale organisms. Given *Sufficiently natural* when we fill in the physical correlate  $\Phi$  in *FLOP schema* ( $(\forall x)(\Phi x \rightarrow \Psi x)$ ), we might get properties fit for particles or the cosmos, but—by *Human disunity*—we will not get properties correlated with anything like a human mental state, and so *Humanly correlated* is lost.

Or starting with *Humanly correlated*, what is wanted is to take the consequents of *FLOPs* to cover the mental states of middling scale organisms. But then, given *Human disunity*, we will not get any sufficiently natural physical correlate and so *Sufficiently natural* is lost.



So I conclude that the physical and mental sides of *FLOPs* are not properly aligned:<sup>13</sup>



<sup>13</sup> Do not be alarmed! This is not a photo of an actual bridge but a computer generated image. Source: <http://www.hoaxorfact.com/Pranks/photographs-showing-misaligned-bridge-constructions-facts.html>

That said, I see two ways out of the clash. The first—which strikes me as gimmicky—splits *FLOP realizers-pain* into many laws, one for each physical realizer:

*FLOP realizer1-pain*: [Fund-law]  $(\forall x) (\text{Fund-phys1 } x \rightarrow \text{Pain } x)$

*FLOP realizer2-pain*: [Fund-law]  $(\forall x) (\text{Fund-phys2 } x \rightarrow \text{Pain } x)$

*FLOP realizer3-pain*: [Fund-law]  $(\forall x) (\text{Fund-phys3 } x \rightarrow \text{Pain } x)$

...

The good news is that the physical correlate is now sufficiently natural in every single law, without any higher level properties or disjunctive properties. But the first piece of bad news is that this proposal explodes the total system of laws (so rejecting the scientific image of a compact totality of fundamental laws: §3.1). And the second piece of bad news is that it splinters the explanatory generalizations, treating states like Fund-phys1 and Fund-phys2 as if they were unrelated antecedents of distinct fundamental laws, having nothing in common but the contingency of a common effect.

The second—and to my mind more serious—way out of the clash involves two synchronized forays into the wild. First, one might (re-)conceive fundamental physics not in terms of material particles but in terms of *information*. In this vein, Wheeler (1990: 5) advances the radical “it from bit” program on which: “[A]ll things physical are information-theoretic in origin,” where:

[E]very item of the physical world has at bottom... an immaterial source and explanation; that which we call reality arises in the last analysis from the posing of yes-no questions and the registering of equipment-evoked responses;...

The intention in adopting informational physics is to allow informational properties to appear in the “*Physical correlate*” slot, as satisfying *Sufficiently natural*.

Secondly, one can adopt a corresponding (re-)conception of the correlates of mental states in informational terms (with the intention of setting up bridge laws from information to the correlates of information, so bringing the physical correlate and the mental state into the right alignment to abut a fundamental bridge law). In this vein, Tononi (2008; see also Balduzzi & Tononi 2009 and Oizumi, Albantakis & Tononi 2014) advances “integrated information theory” (IIT), on which the quantity of consciousness in a system is correlated with its informational integration (“ $\Phi$ ”), and the quality of consciousness is correlated with a further informational quantity (“Q-shape”), such that Balduzzi & Tononi (2009: 21) claim: “[T]he contribution of different brain areas to experience would be mediated (and explained) by how their connectivity, together with their activity patterns, specifies shapes in qualia space.”

Various options in the vicinity of this combination of views are tentatively explored by Chalmers (1996: ch. 8), who think that the correlations arise at the functional level but (1996: 275) recognizes—in accord with the problem of the physical correlate—that “It would be odd if the universe had fundamental laws connecting complex functional organizations to conscious experiences. Rather, one would expect it to be a consequence of simpler, more fundamental psychophysical laws.” Instead Chalmers (1996: 287) suggests:

What we need now is a construct to connect the [basic physical and phenomenal properties]. Information seems to be a simple and straightforward construct that is well suited for this sort of connection, and which may hold the promise of yielding a set of laws that are simple and comprehensive. If such a set of laws could be achieved, then we might truly have a fundamental theory of consciousness.

He (1996: 302–03) connects this with Wheeler’s “it from bit” program, and in later work he (Chalmers & McQueen *forthcoming*: §4) uses IIT as a worked example, for reasons including that it is “mathematically precise” and “offering basic and universal principles connecting consciousness to physical processes.”

The first problem with this combination of informational physics with IIT is just how wild it gets, adopting speculative views of both physics and consciousness, with an attendant price in plausibility. Actually the wildness goes further, for IIT has some panpsychist affinities, in entailing surprising ascriptions of consciousness to simple systems such as thermostats and photodiodes (Tononi & Koch 2015: 11–13). Again all of these speculations may be true in the end. I am just saying that this package is not so plausible for now.

Wildness aside, the more relevant problem with combining informational physics with IIT is that the problem of the physical correlate simply recurs. In informational physics, the unified ground level informational properties are *bit activations*. A message is then a sequence of bit activations. For instance, the message ‘0110100’ is built out of seven binary bits, activated and sequenced in a pattern. Thus Wheeler (1990: 11) speaks of “the bit of information” as being the basic entity posited in informational physics.<sup>14</sup>

In IIT, measures like  $\Phi$  and Q-shape are not bit activations or any other basic informational entities but rather complex, abstractly defined, and multiply realizable informational properties, generated only in very indirect ways from the underlying physics (whether that underlying physics is informational or material). Hence the IIT properties are still not fit for fundamental laws even on an informational conception of the fundamental. Here it is worth recalling Seager’s (1995: 272) complaint against functionalist *FLOPs*:

No other fundamental feature of the world has this character, or a character even remotely like it. It is rather as if one declared that ‘being a telephone’ was a fundamental feature of the world, generated by a variety of physical systems agreeing only in fulfilling the relevant, highly abstract, behaviourally defined functional descriptions.

So suppose that the Q-shape associated with pain is a pyramid, and consider the following toy example of a *FLOP* in IIT-terms:

*FLOP information-pair*: [F-law]  $(\forall x)$  (Pyramidal Q-shape  $x \rightarrow$  Pain  $x$ )

I am saying that “Pyramidal Q-shape” has no more place on the physical side of a fundamental laws than did “C-fiber firing” or “Nociception,” regardless of whether the physics deals in particles or bits.

Note that I am not objecting to informational physics or to IIT, but merely pointing out that the Q-shapes posited by the latter are still not unified ground level features by the lights of the former. My objection is targeted just to the naturalistic dualist who would combine informational physics and IIT to address the problem of the physical correlate. I am saying that, even given both informational physics and IIT, the physical correlates of our mental states are still not fit to abut a fundamental law.

This concludes my discussion of the problem of the physical correlate. I conclude that the physical correlates for the naturalistic dualist’s posited *FLOPs* are subject to conflicting constraints, needing to be sufficiently natural to be fundamental laws, and yet humanly correlated to explain our experiences, when nothing humanly correlated is sufficiently natural for a fundamental law. We are but middling scale organisms and the human correlates of consciousness have no place in fundamental laws. But there are at least four escape routes for the naturalistic dualist:

- Allow fundamental laws to invoke higher level or widely disjunctive properties (§2.1)
- Adopt something like micropsychism plus combination principles, while explaining how the combination principles do not either re-generate the problem or liberate the physicalist (§2.2)
- Allow the fundamental laws to splinter into one-law-per-realizer (§2.3)

---

<sup>14</sup> For Wheeler a bit is understood as a *yes-no* question that an observer puts to a particle and answers with the help of equipment, such as “Is the photon at such a position?”

- Adopt something like informational physics plus an integrated information theory approach to consciousness, while explaining how the problem does not simply re-arise between bits and informational correlates of consciousness such as Q-shape (§2.3)

Each escape represents a foray into the wild, and faces further snares and tangles, but perhaps is still worth the adventure, and perhaps there are other escapes I have missed. But for now I think it fair to conclude that there is a serious problem for the naturalistic dualist.

### 3. Concluding Reflections

In what remains I situate the problem of the physical correlate in three main ways. First, I distinguish it from other known objections to naturalistic dualism. Second, I consider its implications for dualism more generally. Third and finally, I argue that ground physicalism has an advantage here—though it faces problems too.

#### 3.1 Problems for naturalistic dualism

There is a small literature on problems for naturalistic dualism, involving Smart (1959), Seager (1995), Latham (2000), Bourget (2020), and Bennett (*forthcoming*), *inter alia*. Something like the problem of the physical correlate is anticipated in Smart (1959: 143; *epigraph of this paper*). And a range of other problems have been mentioned, but these are not always properly distinguished, validly formulated, or systematically considered.

The problem of the physical correlate ought to be distinguished from at least two other more widely discussed problems for naturalistic dualism. One such problem is *the t-shirt problem*, which is perhaps the main focus of the critical literature, discussed by Adams (1987: 256–58), Latham (2000: 77–80), Bourget (2020: 173–74), and Bennett (*forthcoming*), and recognized by Chalmers (1996: 214) as follows:

Physicists seek a set of basic laws simple enough that one might write them on the front on a T-shirt; in a theory of consciousness, we should expect the same thing. In both cases, we are questing for the basic structure of the universe, and we have good reason to believe that the basic structure has a remarkable simplicity.

The problem is that we have no known way of compactly codifying the psychophysical correlations. We would need both a compact way of thinking of phenomenal states (e.g. position in a multi-dimensional phenomenal space), and a compact mapping of physical correlates into phenomenal space. Resolving just the former issue requires thinking not only that each mode of phenomenal experience (taste, smell, etc.) has its own associated space, but that these spaces can be cobbled into a common phenomenal space. Adams (1987: 256–57) expresses powerful doubts on this score:

[I]s there a unique objectively valid spectrum in which all phenomenal qualia are ordered? Or at any rate a unique phenomenally natural order in which the taste of anise, perhaps, comes between blue and the smell of hydrogen sulfide? Surely not... The different sorts of phenomenal qualia are too diverse from each other for that.

And even if there is a common phenomenal space, we would need to be able to theorize about this space in a simple way with just a few parameters. At this point Latham (2000: 80) objects that:

[T]here would still be an enormous number of fundamental phenomenal parameters, and hence effectively an enormous number of individual laws. And the antecedents of these laws, if physical or functional, would very likely contain a great many variables... This view certainly fails the T-shirt test.

(And all of this is just to get to a common and simple phenomenal space—we still need to map the physical correlates into this space. This all looks daunting.)

More carefully, we might formulate the t-shirt problem as follows:

7. The psychophysical correlations permit no compact codification
8. If the psychophysical correlations permit no compact codification, then the psychophysical correlations are not fundamental laws
9. The psychophysical correlations are not fundamental laws

Here 8 stems from an independently plausible constraint on the total system of fundamental laws, and 7 stems from the daunting challenges that codification faces, involving (i) associating each phenomenal mode with its own space, (ii) cobbling these into a total phenomenal space, (iii) representing position in total phenomenal space with just a few parameters, and (iv) displaying a simple and natural mapping from the space of physical correlates into total phenomenal space. The notion of Q-shape in IIT (§2.3) provides perhaps the best going hope of rebutting 7.

For present purposes I merely wish to point out that the problem of the physical correlate is largely distinct from the t-shirt problem. The problem of the physical correlate arises for *any individual FLOP* (such as *FLOP neural-pain*), while the t-shirt problem rather arises for *the totality of FLOPs*. The potential escapes are different as well, with the naturalistic dualist's primary escape from the t-shirt problem being via IIT, which still does not resolve her problem of the physical correlate (§2.3). The primary point of interaction between these problems is just that one escape that the naturalistic dualist has to the problem of the physical correlate is that of splintering the laws (§2.3), only to be shredded by the t-shirt problem.

A second distinct and better known problem for naturalistic dualism is *the problem of danglers*, which arises in Feigl (1958: 428):

These correlation laws are utterly different from any other laws of (physical<sub>2</sub>) science in that, first, they are nomological “danglers,” i.e., relations which connect intersubjectively confirmable events with events which *ex hypothesi* are in principle not intersubjectively and independently confirmable.

The problem here is that merely tacking *FLOPs* onto the physical laws leave the mental side dangling, in the sense of being a nomic output that is not an input to anything further in the nomic machinery. We might formulate the problem of danglers as follows:

10. The fundamental laws form an integrated causal system
11. If the psychophysical correlations are fundamental laws, then the fundamental laws do not form an integrated causal system
12. The psychophysical correlations are not fundamental laws

The naturalistic dualist could reject 10 and embrace a kind of *epiphenomenalism* for the mental, or could reject 11 by adding on further laws which take mental states as input, such as in *interactionist* proposals on which consciousness can collapse the quantum wave function (Wigner 1967, Chalmers & McQueen *forthcoming*).

For present purposes I merely wish to point out that the problem of danglers is wholly distinct from the problem of the physical correlate (and also from the t-shirt problem). The problem of the physical correlate arises for *any individual FLOP* (such as *FLOP neural-pain*), while the problem of danglers rather arises for *the total system of fundamental laws*. The potential escapes are different as well, with the naturalistic dualist's primary escape from the problem of danglers being epiphenomenalism or interactionism. Smart (1959: 142–43, also 156) distinguishes the problem of the physical correlate and the problem of danglers, saying that (1959: 142) he is both “unable to believe in the nomological danglers themselves” “or in the laws whereby they would dangle.” I am just clarifying that these are distinct concerns.

### 3.2 Implications for dualism

Naturalistic dualism is of course not the only possible form of dualism. The main alternative would be what Bourget (2020) calls “anomalous dualism” which keeps the dualism but drops the *FLOPs*. Indeed Bourget combines explanatory gap arguments with the t-shirt problem (§3.1) to motivate just such a view.

But without *FLOPs*, the dualist looks to have no remaining means to explain the correlations. These correlations hardly look like vast cosmic accidents or pure coincidence. They rather *cry out* for an explanation. From a dualist perspective, the physical and the mental are distinct fundamental elements of reality. Something must be said about why they always pattern together in these ways. But now the dualist seems to be facing explanatory failure: *FLOPs* will not turn the trick (or so I have argued), and—given that we are looking at distinct fundamental properties—nothing else seems open. It seems *purely coincidental*—without any nomological or metaphysical or other reason—that these disconnected fundamental properties correlate at all. Or course it is logically possible that the correlations could be due to some sort of *preestablished harmony*. Perhaps God initiated separate physical and mental sides of reality, and codified their separate laws, just so that these sides of reality would dance together like syncopated wind-up toys. But it is hard to take such options as scientifically serious.<sup>15</sup>

We can regiment the resulting argument against dualism—call it *the correlation problem*—as follows:

13. A viable account of the place of mind in nature must explain the psychophysical correlations
14. Dualism does not explain the psychophysical correlations
15. Dualism is not a viable account of the place of mind in nature

The motivation for 13 is that the correlations hardly look like cosmic accidents, and the motivation for 14 is that *FLOPs* fail (given the problem of the physical correlate: §2, or the t-shirt problem: §3.1, or the problem of danglers: §3.1) but nothing else seems open to the dualist.

So an underlying challenge emerges to dualism generally: either fix up *FLOPs*, or provide some other explanation for the correlations. Perhaps one of these options can be made to work, but for now I conclude that dualism looks like an explanatory failure. I am not saying that the dualist is failing to explain why there are mental states (for she takes that as primitive), but rather that she is failing to explain why our mental states twist in the neural flux.

### 3.3 From naturalistic dualism to ground physicalism (and beyond?)

I have argued that naturalistic dualism faces the problem of the physical correlate, and that dualism generally faces an explanatory failure. I close with a brief sketch as to why such considerations lead towards *ground physicalism* (Schaffer 2009, 2017a, *forthcoming*, see also Rosen 2010 and Bennett 2011, *inter alia*), on which the correlations are understood via vertical ladders of grounding instead of horizontal bridges of causation. As Bennett (2011: 33) says, “physicalism is basically the claim that the physical facts ground the mental facts.” Though I also want to acknowledge some partly analogous problems arising for ground physicalism.

The core difference between naturalistic dualism and ground physicalism is that, where the naturalistic dualist sees physical and mental properties as distinct fundamental properties that are merely nomologically correlated, the ground physicalist thinks of the mental properties as not fundamental but instead grounded in the physical. This encompasses at least four connected differences, concerning (i) the

---

<sup>15</sup> Bourget (2020: 173–78) acknowledges that there are such correlations, and offers “the random theory” to illustrate a potential explanatory option still open. According to this theory, panpsychism obtains plus there just happens to be a kind of harmony between (i) the separate physical and (proto-)phenomenal properties, which just happen to be correlated at the initial conditions; and (ii) the separate physical and phenomenal dynamics, which just happen to run in parallel lines sustaining any correlations. Leaving aside the many wild aspects of the view, I grant that it reduces the explanatory problem to one of explaining the correlations in the initial conditions and dynamical principles, but object that it still fails to explain either of these matters.

source of the psychophysical correlations, (ii) their modal strength, (iii) the fundamentality status of the properties involved, and (iv) the extent to which the properties count as separate and distinct:

	<i>Naturalistic dualism</i>	<i>Ground physicalism</i>
<i>Source</i>	Law of nature	Grounding principle
<i>Strength</i>	Nomological necessity	Metaphysical necessity
<i>Status</i>	Both fundamental	Physical fundamental
<i>Separation</i>	Distinct	Not distinct but connected

(I assume that the reader is familiar enough with this ground-theoretic ideology. For further clarification see Schaffer 2009, 2016, and *forthcoming*.)

These options are exclusive but not exhaustive. For instance, there is also the option of maintaining a stronger version of physicalism involving a reductive identity. For present purposes I am just focusing on the comparison between naturalistic dualism and ground physicalism, so will leave other options such as reductive physicalism off the table. Though the problems faced by both naturalistic dualism and ground physicalism may help motivate reductive views (I return to this thought at the close).

Focusing just on the comparison between naturalistic dualism and ground physicalism, my initial thought (which started me towards this paper) was that the problem of the physical correlate serves to favor ground physicalism, fitting a perspective on which the psychophysical correlations are understood via vertical ladders of grounding instead of horizontal bridges of causation. For let us compare paradigmatic examples of fundamental laws like  $F=ma$ , to paradigmatic examples of inter-level grounding such as seen in vertical connections from the chemical up to the biological. The fundamental laws of our world seem to concern the global dynamics of natural properties:

- They link sufficiently natural properties (no higher level or widely disjunctive properties)
- They are global, concerning whole states of the cosmos
- They are dynamic, governing temporal evolution (or expressing global constraints)

In this vein Maudlin (2017: 172) speaks of “fundamental laws of temporal evolution” which “specify how the state of the universe will, or might, evolve from a given initial state.”<sup>16</sup>

In contrast, inter-level grounding connections seem to concern local synchronic links from more to less natural properties:

- They do not link sufficiently natural properties, but rather involve (progressively) higher level properties linking relatively lower level properties relatively higher level properties
- They are regional, covering middling scale systems
- They are synchronic, concerning how the relatively lower level properties of a system at a time fix its higher level properties at that time

---

<sup>16</sup> This perspective on fundamental laws goes back to Russell. As Loewer (2007: 296) summarizes: Russell (1913) observed that the fundamental laws—he was thinking of the differential equations of classical mechanics, but the same holds for quantum mechanics—specify how the whole state of an isolated system evolves (or the chances of possible evolutions) but don’t specify which parts of the state at one time are causally connected to which parts of the state at other times.



For instance, to the extent that my various chemical properties fix my biological property of *being alive*, this is a vertical link through higher level properties, covering a middle scale organism, at a single fixed time.

The psychophysical correlations are regional synchronic correlations involving higher level properties. So the psychophysical correlations seem to fit the image of inter-level grounding connections and not the image of fundamental laws, at least in the three respects bulleted. Advantage: ground physicalism.

Moreover, one can see how the problem of danglers (§3.1) immediately disappears for the ground physicalist. The psychophysical connections are not added fundamental laws of nature, so no dangling mental output or other complications might compromise the integrity of the existing physical machinery. Does any analogous issue arise? It is not clear that there is any analogous integrity requirement on the grounding machinery, but at any rate mental states for the ground physicalist are well-integrated into the grounding story, being grounded from relatively lower level neurobiological states, and in turn grounding relatively higher level sociological states (given the plausible assumption that the psychological at least partly grounds the sociological), just as on a classical physicalist “levels of nature” image (as in Oppenheim & Putnam 1958).

Connecting back to the problem of the physical correlate (§2), ground physicalism readily resolves the problem as stated simply because it does not posit any new fundamental laws. As to the underlying dilemma between satisfying *Sufficiently natural* and *Humanly correlated*, it was only through fundamental laws that *Sufficiently natural* was motivated (§2.1). Grounding links are not constrained to satisfy *Sufficiently natural*, but can involve higher level properties on either side. So we can think of the psychophysical correlations—in accord with *Humanly correlated*—in terms of the grounding links from the neural to the psychological. Neural properties are too high level for fundamental laws, but not for the grounding links from the neural to the psychological. Just as the chemical to biological links involve middling level chemical properties on the left, so the neural to psychological links involve middling level neural properties on the left.

But there is a partly analogous problem arising, at least given a further perspective I endorse, on which grounding links are mediated by “grounding principles” or “laws of metaphysics” (Schaffer 2017b), with *root principles* of grounding akin to fundamental laws of nature, assessed on a holistic abductive basis for overall simplicity and strength (Schaffer *forthcoming*: §1.3). For now the ground physicalist must state the root principles through which the neural grounds the psychological. And it seems implausible that anything as specific as neural properties should show up in the root principles of vertical generation. (I do not say that it is impossible for nature to have such root principles, but I still grant that it seems implausible.)

What is the relevant plausibility constraints on the root principles? My thought is that something important shifts, in terms of constraints on fundamental laws of nature versus root principles of grounding. For *Sufficiently natural* is appropriate for fundamental laws (§2.1) but not plausible for root principles, since these are the underlying vertical connections and so need to span levels. Rather my suggestion is that a better constraint on root principles is something like *Topic neutrality*, in the sense that the root principles should not invoke level-specific properties but rather general level-neutral properties such as causal and spatiotemporal properties. (Just as the root principles of set formation and mereological fusion work across their respective hierarchies, without consideration as to the stages of their inputs, so I am thinking of the root principles of level generation as working across the levels without consideration of specific levels.)

The classic form of *Topic neutrality* satisfying claims is *functionalist*. So I think this issue may drive the ground physicalist to a general functionalist stance on all higher level structure, and so to a view of psychological structure that I elsewhere (Schaffer *forthcoming*) call *Ground functionalism*. But the important point here, vis-a-vis the problem of the physical correlate, is that *Topic neutrality* and *Humanly correlated* are not in conflict, at least provided that there is some (functionalist or otherwise) way of characterizing human neural and mental states in topic neutral terms so that their grounding connection might be subsumed under a general and neutral root principle.

Concluding with the t-shirt problem (§3.1), ground physicalism readily resolve the problem as stated simply because it does not posit any new fundamental laws. But a partly analogous problem arises, at least given the perspective on grounding links as involving root principles. Insofar as posited root principles are assessed on a holistic abductive basis for overall simplicity and strength, they need to be kept simple, and the worry remains that no compact statement of the phenomenal connections is possible.

I think that the t-shirt remains a serious problem even for ground physicalism. As mentioned above, the notion of Q-shape in IIT (§2.3) provides perhaps the best going hope of resolving the problem, but the posited IIT connections are not functionalist or otherwise topic neutral. Perhaps there is a way to reconcile IIT and topic neutrality, but this needs to be shown, and so an underlying worry emerges that the ground physicalist may be driven to *incompatible solutions* for her analogue of the problem of the physical correlate (functionalism) and her analogue of the t-shirt problem (IIT). More work is needed.

So I see the t-shirt problem as a residual shared difficulty for both naturalistic dualism and ground physicalism. Reductive physicalism may have the advantage here, insofar as there can just be various identities strewn across nature, without need for compactification.

I conclude that ground physicalism is in a better position than naturalistic dualism with respect to the problem of the physical correlate, but there are still problems arising.<sup>17</sup>

## References

- Albahari, Miri 2020. Beyond Cosmopsychism and the Great I Am: How the World might be Grounded in Universal (Advaitic) Consciousness. *The Routledge Handbook of Panpsychism*, ed. William Seager: 119–30. Routledge.
- Adams, Robert 1987. Flavors, Colors, and God. *The Virtue of Faith and Other Essays in Philosophical Theology*: 243–62. Oxford University Press.
- Armstrong, D. M. 1978. *Universals and Scientific Realism, Volumes 1–2*. Cambridge University Press.
- 1983. *What is a Law of Nature?* Cambridge University Press.
- Balduzzi, David and Giulio Tononi 2009. Qualia: The Geometry of Integrated Information. *PLoS Computational Biology* 5: e1000462.
- Bell, J. S. 1990. Against ‘Measurement’. *Physics World* 3: 33–40.
- Bennett, Karen 2011. By Our Bootstraps. *Philosophical Perspectives* 25: 27–41.
- *forthcoming*. Why I am Not a Dualist. *Oxford Studies in Philosophy of Mind* 1.
- Bourget, David 2020. Anomalous Dualism: A New Approach to the Mind-Body Problem. *The Routledge Handbook of Panpsychism*, ed. William Seager: 168–80. Routledge.
- Chalmers, David 1995. Facing up to the Problem of Consciousness. *Journal of Consciousness Studies* 2: 200–19.
- 1996. *The Conscious Mind: In Search of a Fundamental Theory*. Oxford University Press.
- 2003. Consciousness and its Place in Nature. *The Blackwell Guide to Philosophy of Mind*, eds. Stephen Stich and Ted Warfield: 102–42. Blackwell.
- 2016. The Combination Problem for Panpsychism. *Panpsychism: Contemporary Perspectives*, eds. Godehard Brüntrup and Ludwig Jaskolla: 179–214. Oxford University Press.
- Chalmers, David and Kelvin McQueen *forthcoming*. Consciousness and the Collapse of the Wave Function. *Consciousness and Quantum Mechanics*, ed. Shan Gao. Oxford University Press.
- Crick, Francis and Christof Koch 1998. Consciousness and Neuroscience. *Cerebral Cortex* 8: 97–107.
- Feigl, Herbert 1958. The ‘Mental’ and the ‘Physical’. *Minnesota Studies in the Philosophy of Science* 2: 370–479.
- Garcia-Larrea, Luis and Hélène Bastuji 2018. Pain and Consciousness. *Progress in Neuro-Pharmacology and Biological Psychiatry* 87B: 193–99.
- Gertler, Brie *forthcoming*. Dualism: How Epistemic Issues Drive Debates about the Ontology of Consciousness. *Oxford Handbook of the Philosophy of Consciousness*. Oxford University Press.
- Goff, Philip 2009. Why panpsychism doesn’t help us explain consciousness. *Dialectica* 63: 289–311.

<sup>17</sup> Thanks to David Chalmers, Uriah Kriegel, and Adam Pautz.

- . 2017. *Consciousness and Fundamental Reality*. Oxford University Press.
- Hicks, Michael Townsen and Jonathan Schaffer 2017. Derivative Properties in Fundamental Laws. *British Journal for the Philosophy of Science* 68: 411–50.
- Kim, Jaegwon 1998. *Mind in a Physical World: An Essay on the Mind-Body Problem and Mental Causation*. The MIT Press.
- Kriegel, Uriah *forthcoming*. What is the Philosophy of Consciousness? *Oxford Handbook of the Philosophy of Consciousness*. Oxford University Press.
- Latham, Noa 2000. Chalmers on the Addition of Consciousness to the Physical World. *Philosophical Studies* 98: 71–97.
- Lewis, David 1983. New Work for a Theory of Universals. *Australasian Journal of Philosophy* 61: 343–77.
- Loewer, Barry 2007. Laws and Natural Properties. *Philosophical Topics* 35: 313–28.
- Maudlin, Tim 2007. *The Metaphysics within Physics*. Oxford University Press.
- Melzack, Ronald 1999. From the GATE to the Neuromatrix. *Pain* S6: S121–26.
- Melzack, Ronald and Patrick Wall 1965. Pain Mechanisms: A New Theory. *Science* 150: 971–89.
- Mill, J. S. 1843. *A System of Logic, Ratiocinative and Inductive, Being a Connected View of the Principles of Evidence, and the Methods of Scientific Investigation, vol. 2*. John W. Parker.
- Nagasawa, Yujin and Khai Wager 2016. Panpsychism and Priority Cosmopsychism. *Panpsychism*, ed. Godehard Bruntrup and Ludwig Jaskolla: 113–29. Oxford University Press.
- Nida-Rümelin, Martine 2007. Dualist Emergentism. *Contemporary Debates in Philosophy of Mind*, eds. Brian McLaughlin and Jonathan Cohen: 269–86. Blackwell.
- North, Jill 2013. The Structure of a Quantum World. *The Wave Function: Essays on the Metaphysics of Quantum Mechanics*, eds. Alyssa Ney and David Albert: 184–202. Oxford University Press.
- Oizumi, Masafumi, Larissa Albantakis, and Giulio Tononi 2014. From the Phenomenology to the Mechanisms of Consciousness: Integrated Information Theory 3.0. *PLoS Computational Biology* 10: e1003588.
- Oppenheim, Paul and Hilary Putnam 1958. Unity of Science as a Working Hypothesis. *Minnesota Studies in the Philosophy of Science* 2: 3–36.
- Pautz, Adam 2009. A Simple View of Consciousness. *The Waning of Materialism*, eds. Robert Koons and George Bealer: 25–66. Oxford University Press.
- . 2017. The Significance Argument for the Irreducibility of Consciousness. *Philosophical Perspectives* 31: 349–407.
- Putnam, Hilary 1975. Philosophy and our Mental Life. *Philosophical Papers, Volume 2*: 291–303. Cambridge University Press.
- Rabin, Gabriel 2019. Grounding the Gaps or Bumping the Rug? On Explanatory Gaps and Metaphysical Methodology. *Journal of Consciousness Studies* 26: 191–203.
- Rosen, Gideon 2010. Metaphysical Dependence: Grounding and Reduction. *Modality: Metaphysics, Logic, and Epistemology*, eds. Bob Hale and Aviv Hoffmann: 109–36. Oxford: Oxford University Press.
- Schaffer, Jonathan 2009. On What Grounds What. *Metametaphysics*, eds. David Chalmers, David Manley, and Ryan Wasserman: 347–83. Oxford University Press.
- . 2016. Grounding in the Image of Causation. *Philosophical Studies* 173: 49–100.
- . 2017a. The Ground between the Gaps. *Philosophers' Imprint* 17: 1–26.
- . 2017b. Laws for Metaphysical Explanation. *Philosophical Issues* 27: 302–21.
- . *forthcoming*. Ground Functionalism. *Oxford Studies in Philosophy of Mind* 1.
- Seager, William 1995. Consciousness, Information, and Panpsychism. *Journal of Consciousness Studies* 2: 272–88.
- Shani, Itay 2015. Cosmopsychism: A Holistic Approach to the Metaphysics of Experience. *Philosophical Papers* 44: 389–437.
- Smart, J. J. C. 1959. Sensations and Brain Processes. *The Philosophical Review* 68: 141–56.
- Stoljar, Daniel 2006. *Ignorance and Imagination: The Epistemic Origin of the Problem of Consciousness*. Oxford University Press.
- Tononi, Giulio 2008. Consciousness as Integrated Information: A Provisional Manifesto. *The Biological Bulletin* 215: 216–42.
- Tononi, Giulio and Christof Koch 2015. Consciousness: Here, There, and Everywhere? *Philosophical Transactions of the Royal Society B* 370: 20140167.

- von Neumann, John 1955. *Mathematical Foundations of Quantum Mechanics*, trans. Robert Beyer. Princeton University Press.
- Wheeler, John Archibald 1990. Information, Physics, Quantum: The Search for Links. *Complexity, Entropy, and the Physics of Information*, ed. Wojciech Zurek: 3–28. Addison-Wesley.
- Wigner, E. P. 1967. Two Kinds of Reality. *Symmetries and Reflections: Scientific Essays of Eugene P. Wigner*, eds. Walter Moore and Michael Scriven: 185–99. Indiana University Press.