Anomalous Dualism: A New Approach to the Mind-Body Problem*

David Bourget

University of Western Ontario

July 26, 2017

A satisfactory solution to the mind-problem should answer the two following questions: i) are phenomenal properties, the properties that characterize states of consciousness, physical? ii) how do phenomenal properties causally interact with physical properties? To a first approximation, physicalism and dualism are the two possible answers to the first question. There are three kinds of views regarding causal interactions between phenomenal and physical properties: nomism (they interact through deterministic laws), acausalism (they do not causally interact), and anomalism (they interact, but not through deterministic laws). In this paper, I explore anomalous dualism, a combination of views that has not previously been explored. I suggest that a kind of anomalous dualism, nonreductive anomalous panpsychism,

*For some time, this paper was circulated under the title "Anomalous Panpsychism".

promises to offer the best overall answer to two pressing issues for dualism: the problem of mental causation and the mapping problem (the problem of predicting mind-body associations). I will start by charting the logical space around anomalous dualism.

1 The logical space

We can define *physical properties* as properties that satisfy one of the following conditions: i) properties of roughly the kind that have so far been uncovered by biology, chemistry, and physics; ii) the properties that can be constituted by physical properties; iii) the properties that actually constitute physical properties (if any). By X *constitutes* Y, I mean that X realizes Y, grounds Y, or stands in a determinate-determinable relationship to Y. This definition of physical properties corresponds roughly to what Chalmers (2015) calls "broadly physical properties".

I take *physicalism* to be the view that everything, including phenomenal properties, is physical. *Anti-physicalism* is the negation of physicalism, whereas *dualism* is the view that phenomenal properties are instantiated but none is physical. Anti-physicalism without dualism is quite implausible, so the mind-body problem is typically conceived of as requiring us to choose between physicalism and dualism. Note that, since we understand "physical" broadly, Russellian monism and functionalism both count as kinds of physicalism.

Issues pertaining to mental causation figure centrally in the debate between physicalists and dualists. For present purposes, we can think of causation as a relation between events, and we can think of events as instantiations of properties by individuals at times. There are three possible views that one might take on any given alleged instance of causation between mental and physical events: *nomism* is the view that the instance of causation is subsumed under a deterministic law of nature; *anomalism* is the view that it is an instance of non-deterministic causation; *acausalism* is the view that denies that there is any genuine causation. One can be a nomist (or an anomalist or an acausalist) about some alleged interactions and not others, but, plausibly, mind-to-matter interactions are either all nomic or all anomalous, and matter-to-mind interactions are either all nomic or all anomalous. I will refer to nomism about all mental-physical causation in both directions as *generalized nomism*, and likewise for the other views.

Physicalism and dualism can both be combined with each of nomism, anomalism, or acausalism about any given subset of alleged causal relations. If we consider only generalized versions of the views on causation and set aside acausalism as implausible, we have four options shown in Table 1. Only one

Table 1: Four possible positions on the mind-body problem

	Generalized nomism	Generalized anomalism
Physicalism	Standard physicalism	Anomalous monism
Dualism	Standard interactionism	Anomalous dualism

of these views has not been seriously considered: anomalous dualism. The

aim of this paper is to explore the viability of this view as an approach to the mind-body problem. More specifically, I am interested in finding out whether this view is promising *for a dualist view*.

For someone who is persuaded that physicalism is true, this might seem like a futile exercise, but I think there is something in this paper even for convinced physicalists. Physicalism is largely motivated in opposition to dualism: dualism has too many problems (especially pertaining to mental causation), so physicalism must be true despite appearances (no one can deny that phenomenal properties really don't seem to be physical). If anomalous dualism turns out to avoid dualism's well-known difficulties, the motivation for physicalism will be weakened.

In the next two sections, I consider the two main problems that dualism faces: the problem of mental causation and the mapping problem. The latter has not been widely discussed, but it seems to me that it underlies much of the skepticism about dualism. I will suggest that anomalous dualism is the most promising dualist view as far as mental causation and the mapping problem are concerned. This is not to say that anomalous dualism does not have other problems. In section 4, I will consider various potential issues with anomalous dualism. I will suggest that a version of anomalous dualism that is also a version of panpsychism could conceivably solve many of anomalous dualism's apparent problems.

2 Mental efficacy

Dualism has troubles with mental causation. The problem is that the physical world appears to be closed to causal influence from non-physical factors, in the sense that nothing non-physical can be causally relevant to physical events. If this is true, then either dualism is false or mental events are not causally relevant to physical events. Assuming that mental events are causally relevant to physical events, dualism must be false.

In order to see possible ways out of this argument, we need to get a little clearer on the justification for the claim that nothing non-physical can influence physical events. This does not seem to be something that is revealed by physics or any other science.[1] To a first approximation, what physics seems to tell us is that every physical event is a deterministic effect of another physical event. This is what is generally referred to as the *completeness* of physics. For reasons that will become apparent shortly, I prefer to refer to this thesis as *physical determinism*.

Physical determinism Every physical event is an immediate deterministic effect of a physical event.

By *immediate*, I mean that the effect is not mediated by another event.

It is not entirely obvious how to use physical determinism to argue against dualism, but there is at least one widely accepted way. Suppose first that

[1] Papineau (2001) makes this point and offers an excellent overview of the history of the closure problem for dualism. The first argument discussed below is adapted from Papineau's discussion.

mental events have physical effects.

Mental efficacy Some phenomenal events have physical effects.

To complete the argument, we need to assume that the effects of mental events on physical events are not all overdetermined.

No systematic overdetermination The immediate effects of phenomenal events on physical events are not all overdetermined.

An event is *overdetermined* when there are two wholly distinct events of which it is an immediate effect. By *wholly distinct*, I mean events that are non-identical and not related through a relation of constitution (as defined above).[2]

Given mental efficacy and the no systematic overdetermination principle, we can infer that some phenomenal event M has an immediate physical effect P that is not overdetermined. P's not being overdetermined means that there is not a second event wholly distinct from M that also has P as immediate effect. But physical determinism requires that P is the immediate effect of a physical cause C. So M must not be wholly distinct from C. Given our definition of "physical", this means that M is a physical event. Since we are assuming that events are instantiations of properties, it follows that some mental properties are physical properties. Therefore, dualism is false.

Since the no systematic overdetermination principle is not a priori, one might question it. The rationale behind it is that it is hard to see what

[2]The "wholly distinct" qualification makes no-overdetermination consistent with nonreductive physicalism and Russellian monism.

plausible setup might guarantee the overdetermination of the effects of mental events on physical events. Overdetermination can happen, as in firing squad situations, but no one has made it plausible that the effects of mental events on physical events are systematically overdetermined. Without some kind of independent justification, the hypothesis of systematic overdetermination is ad hoc and implausible.

One widely overlooked reason why denying the no systematic overdetermination principle is unattractive is that this premise is actually dispensable. It is dispensable on the assumption that mental efficacy entails counterfactual dependence.

Counterfactual dependence Some actual physical event is counterfactually dependent on an actual mental event.

An event A is *counterfactually dependent* on an event B if and only if it is the case that A would not have occurred if B had not occurred. Even if mental efficacy didn't entail counterfactual dependence, counterfactual dependence is just as plausible as mental efficacy, so we can replace mental efficacy by counterfactual dependence in our argument. This yields the following argument against dualism. Take any physical event P that is counterfactually dependent on a phenomenal event M as required by counterfactual dependence. Physical determinism implies that P had an immediate sufficient physical cause C1, which itself had an immediate sufficient physical cause C2, and so on ad infinitum (or up to a first physical event if we make an exception to physical determinism for a first physical event). Take the event C along

this chain that occurred at exactly the same time as M (we can slice the events of the chain in whatever way is necessary to delineate such an event). If dualism is true, C is wholly distinct from M. It follows from the standard, non-backtracking way of assessing counterfactuals that C would have occurred even if M had not occurred.[3] Since C is by hypothesis sufficient for a chain of events leading to P, this means that P would have occurred even if M had not occurred. This contradicts our assumption that P is counterfactually dependent on M. Therefore, one of physical determinism, counterfactual dependence, or dualism must be false.

Denying counterfactual dependence is an option for the dualist, but it comes with a huge cost: if counterfactual dependence is false, there is a clear sense in which the mental makes no difference to the physical. But if mental states made no difference to the physical, we wouldn't expect brains created through natural selection to involve any consciousness, much less for the phenomenal properties associated with physical properties to exhibit any sort of cohesion or "make sense". This seems to be a fatal objection to mind-to-matter acausalism however the relevant causal roles are understood

[3]In assessing counterfactuals of the form "had A been the case, B would have been the case", we must check whether B is the case at the nearest possible worlds at which A is the case. Importantly, the nearest worlds need not be worlds that are physically possible: we don't look for worlds that have histories that explain A through laws like ours (this would typically require us to "backtrack" into the history of the world). Rather, we allow A to be the case at a world just like ours, without a suitable history, as if by miracle (Lewis 1973, 1979). As a result, any contemporaneous or earlier fact at the actual world that is metaphysically compatible with A is also a fact at the nearest A worlds. In our case C is contemporaneous with M and metaphysically independent of it (in virtue of dualism), so the nearest M-less world is a C world.

exactly.[4]

This leaves a dualist with one possible response to the arguments from mental causation, which is to deny physical determinism. Unlike the other responses we have considered so far, this response has some initial plausibility. After all, physical determinism has been shown to be false by quantum mechanics. More specifically, quantum measurements are indeterministic in two ways: first, the outcomes of quantum measurements are probabilistic, following probability distributions fixed by quantum states; second, when a measurement occurs, and what the measured observable is, are not determined by anything within standard quantum theory.[5] These facts are typically ignored by philosophers of mind because the evidence from neuroscience seems to suggest that mental activity is implemented in macroscopic neural patterns.[6] But our understanding of the brain remains fairly limited, and it is not hard to imagine how sub-neuronal conditions for consciousness might have gone unnoticed. So, let us not prejudge this empirical question.

[4]One might also argue from everyday experience that phenomenal states have bodily effects. For example, conscious intentions seem to cause bodily movements. Some considerations in section 4 below throw some doubt on this, but they leave untouched the argument that consciousness must do something for our brains to have evolved to produce cohesive streams of experience.

[5]There are developments of the theory that attempt to give a physical explanation of measurement, but none is widely accepted.

[6]Sometimes, allowances are made for quantum randomness by formulating completeness as follows: the objective probability of every physical event is determined by a prior physical event (Yablo 1992; Bennett 2007). This is still too strong because *that* a measurement will occur at any given time in a quantum system is not physically determined on standard interpretations. Events that could occur as a result of measurement don't have any objective probability *until* a measurement is "decided". If measurement is not decided by a physical state (which is not the case on standard QM), not every physical event has a probability fixed by a prior physical event.

If we limit ourselves to assuming a limited completeness of physics that is consistent with quantum mechanics, the following is the principle that we should use as premise as part of arguments from mental causation against dualism:

Completeness For every non-random physical event Y, some immediately prior physical event X is causally sufficient for Y.

A *random event* is an event whose occurrence was not determined by deterministic laws of nature, for example, the immediate outcome of a quantum measurement. I formulate the exception to determinism built into completeness in terms of randomness, and not specifically in term of quantum measurement, because I want to stay as close as possible to observation. We don't know whether a future, more complete physics might not drop "measurement" talk entirely in favor of a deeper characterization of what is going on. What we do know is that nature occasionally exhibits fundamentally indeterministic behavior that is correctly modeled by the mathematics of quantum mechanics.

Our qualified, scientifically correct completeness claim renders invalid the arguments sketched above. However, the dualist is not out of trouble yet, because the arguments can be fixed. To make our arguments valid again, we can add a further premise that specifies that the effects of mental events do not fall under the exception we have carved out for random physical events:

No-randomness The immediately physical effects of mental events are not

random events.

This might seem like a somewhat ad hoc claim, but it falls out of a slightly stronger claim that seems to be a natural, widely held view:

Mind-to-matter nomism The effects of mental events on physical events fall under deterministic laws.

Given mental efficacy and no-randomness, we can infer that mental events have immediate, non-random physical effects. By the completeness principle, each of these effects has an immediate physical cause. Assuming that there is no systematic overdetermination, some of the physical causes of the effects of mental events are not wholly distinct from the mental events. This leads to the conclusion that dualism is false.

The argument from counterfactual dependence is slightly more complicated with completeness than with physical determinism. It requires us to consider not just any counterfactual dependence relationship between mental and physical events, but a *direct* counterfactual relationship, one that is not the result of intermediary counterfactual dependence relationships. We can infer the existence of such relationships from counterfactual dependence: if there is a counterfactual dependence relationship between M and P, there must be a direct counterfactual dependence relationship between some mental event N and some physical event O, because the mental and the physical have to interface somewhere along the chain of dependent events. Now suppose that C is the physical event at the origin of the chain of physical events

leading to O, which we know exists by completeness and no-randomness. Either C occurred after N, or it did not. We can plausibly argue that C cannot have occurred after N because this would mean that it occurred in the time between N and O, but there cannot be any such time: if there was, we could discern intermediary events and counterfactual dependence relationships between N and O. If C did not occur after N, the argument proceeds as before: we can find a link in the chain C...O that is contemporaneous with N and argue from dualism that it would have occurred even if N had not occurred, which means that O (and P) would have occurred even if N (and M) had not occurred.

The versions of the argument from mental causation that appeal to completeness don't leave the dualist many options. Completeness is too well established scientifically to be questioned on the basis of a posteriori armchair considerations, and denying mental efficacy altogether seems absurd. As a result, denying mind-to-matter nomism seems to be the only potentially acceptable option. The thesis that mind-to-matter nomism is false has received some attention in the context of Davidson's (1970) defense of anomalous monism, and there seems to be fairly widespread agreement that Davidson at least makes a good case for doubting the existence of strict laws connecting mental and physical events.[7] Many authors have raised doubts regarding the viability of anomalous monism as an account of mental causation, but anomalism remains fairly plausible independently of the rest of

[7]Seager (1981) offers a rigorous defense of anomalism that refines Davidson's arguments.

Davidson's view. In any case, it is far more plausible than denying completeness or denying mental causation altogether.

The dualist should not merely deny mind-to-matter nomism; she should also endorse mind-to-matter anomalism. First, this is required to escape the argument from counterfactual dependence, which only needs one instance of deterministic mind-to-matter causation. Second, if she took the position that some but not all effects of mental events on physical events are non-random, she would run the risk that some of the non-random events are also not overdetermined, which would allow her opponent to use our argument from no systematic overdetermination. Overdetermination plausibly can only occur by accident, so an "almost no overdetermination" principle is about as plausible as our no systematic overdetermination principle. As a result, the dualist has to at least endorse the view that deterministic causation from mental to physical events is rare. She might as well take the simpler, more principled view that it is the nature of such causation to involve some randomness (mind-to-matter anomalism).

Endorsing mind-to-matter anomalism seems to be the best strategy for the dualist, but one might worry that this response is merely an empty shell of a theory, and that it is not plausible unless there is some reasonable way of filling in the details, of explaining how the supposed random interface between the mental and the physical works.

A number of theorists have argued for a relationship between quantum randomness and consciousness. However, I am not aware of a "quantum the-

ory of consciousness" that succeeds at giving a reasonably plausible account of mental causation that is consistent with completeness, mind-to-matter anomalism, and the totality of empirical evidence regarding the dynamics of physical systems. Some theories allow mental states to deterministically cause physical states, claiming that such causation occurs under the cover of quantum randomness (cf. Eccles 1978), which is why we have not noticed it. This kind of view is ruled out by our arguments from completeness, as well as by empirical evidence (Bourget 2004). Other theories are consistent with completeness but not the totality of evidence regarding the dynamics of physical systems. According to Stapp's (1996) view, for example, the conscious mind acts on the body by selecting when a quantum measurement is performed and what observable is measured. Because of a phenomenon known as the "Zeno effect", this theoretically allows the conscious mind to control what the brain does. This is an example of a possible way of fleshing out mind-to-matter anomalism. However, this view turns out to be inconsistent with simple empirical observations (see Bourget 2004).

In order for the "quantum response" to the problem of mental causation for dualism to be plausible, it has to respect the constraint that quantum measurements have very slight effects on macroscopic systems as far as we can observe. To my knowledge, no theory that clearly respects this constraint has been offered. However, it remains that mind-to-matter anomalism seems to be the dualist's best possible solution to the problem of mental causation.

3 The mapping problem

The previous section motivated mind-to-matter anomalism. In this section, I present a motivation for matter-to-mind anomalism. While some authors have considered mind-to-matter anomalism, I don't think that anyone has considered combining dualism with matter-to-mind anomalism. It seems to be almost universally assumed that if the mind is not physical, it arises via deterministic psychophysical laws. In this section, I want to point out that it is not obvious that this assumption is correct.

If dualism and matter-to-mind nomism were both true, the psychophysical laws that govern matter-to-mind causation would be fundamental laws of nature. As Chalmers (1996) points out, such a view would only be plausible if the psychophysical laws could be given a simple, general formulation (roughly, one that fits on a t-shirt). Otherwise, the relevant psychophysical laws would not be plausible candidate fundamental laws of nature. A canonical statement of the psychophysical laws should also explicitly relate mental states and physical states under their mental and physical descriptions, respectively. For example, it would not do to say simply that "a physical state gives rise to the mental state associated with it". A suitable statement of a psychophysical law would have to relate full descriptions of mental states as such with physical states under their full physical descriptions. Call a general, simple statement specifying which mental states (under a mental description) occur in any given physical condition (under a physical description) a *general*

psychophysical mapping. Matter-to-mind nomic dualism seems to require the existence of a general psychophysical mapping, in the sense that the former would be quite implausible if we knew that there is no such mapping to be found.

I want to suggest that we have fairly good evidence that there is no general psychophysical mapping. Note first that the existence of psychophysical correlations does not imply the existence of a general psychophysical mapping. It could be that there is a perfect correlation between phenomenal properties and physical properties, in the sense that the same phenomenal property is always instantiated along with the same physical property and vice versa, yet there is no general psychophysical mapping. Perfect correlation is even consistent with the absence of a finitely statable psychophysical mapping (a minimally demanding understanding of the requirement that a general psychophysical mapping be "simple").

Neuroscience has revealed numerous correlations between brain areas and types of conscious experience and other kinds of mental activity. It has also revealed what appear to be limited mappings between aspects of conscious experience and certain kinds of brain activity. For example, the phenomenological color space can plausibly be mapped in a relatively straightforward way to dimensions of activation in certain neural networks in the brain (see Churchland 1986). These are impressive findings, but they fall far short of a general psychophysical mapping. The associations that we know exist between phenomenal properties and physical properties don't seem to fall under

a broad pattern that suggests a general psychophysical mapping. If we were to plot the known correlations between physical and phenomenal properties, we would see some local patterns (as in the case of color experience), but, aside from these local patterns, the points would jump all over the place, forming no recognizable curve that we can characterize. We have no idea how to extrapolate a general psychophysical mapping from what we have. Such a mapping has not even been *imagined*.[8]

I don't wish to diminish the accomplishments of neuroscience in any way: an extremely impressive number of correlations have been found. On the contrary, my point is that the vast quantity of correlational data that has been collected, together with the absence of a candidate general psychophysical mapping, makes it plausible that no psychophysical mapping is to be found. The more we find out about psychophysical correlations, the more non-generalizable they seem.

One philosophical theory initially seems to help with the mapping problem: representationalism. *Representationalism* is roughly the view that an experience of a quality Q is a mental state that phenomenally represents Q.[9] If we could give an account of the physical basis of phenomenal representation, then it seems that we would have a general psychophysical mapping.

[8]Chalmers' (1996) information-theoretic proposal might be closest that we have come.

[9]For defenses of non-reductive representationalism, see Byrne (2001), Crane (2003), Chalmers (2004), Pautz (2009), Bourget and Mendelovici (2014), and Bourget (forthcoming b). Some representationalists suggest that a further ingredient might be required in addition to phenomenal representation of a content: an intentional mode or representational manner (see Crane 2003, Chalmers 2004, Speaks 2010, 2015). I argue against such extra ingredients in Bourget (2015, 2017, forthcoming a).

Suppose that physical relation R is the physical basis of phenomenal representation. Then it seems that we have the following general psychophysical mapping:

Representationalist mapping An individual experiences quality Q just in case they stand in R to Q.

Suppose, for example, that R is the relation that one stands in to X just in case one has an internal state that is cognitively integrated and "tracks" the presence of X in the world. Call representationalism combined with this account of R *non-reductive tracking representationalism*, or *NRT representationalism*. Given NRT representationalism, it seems that we can predict what sort of experience one has in any given condition. For example, if one stands in R to redness, then one will experience redness.

Despite appearances, NRT representationalism does not provide a general psychophysical mapping, even if R can be specified in a plausible way. A general psychophysical mapping is supposed to relate phenomenal properties *under a phenomenal description* with physical properties *under a physical description*. NRT representationalism does not specify such a mapping. The problem is that "experiencing Q" and "tracking Q" cannot at the same time be phenomenal and physical descriptions of phenomenal states and physical states, respectively, because "Q" is couched either in phenomenal language or in physical language.

To make the problem more vivid, consider how NRT representationalism might try to predict which mental state is associated with any given physical

state. The idea is that we check which property is tracked. If an individual is tracking Q in the right way, we can predict that the individual is experiencing Q. The problem with this is that, when we identify what is tracked, we identify it under a physical description, for example, "the property of reflecting electromagnetic radiation of about 650 nm" (*R650*, for short). "Experiencing R650" is not a phenomenal description of the state of experiencing red—the proper phenomenal description of that state is "experiencing red". So the NRT representationalist account does not make any prediction about experiencing red under a phenomenal description. This fact is easily overlooked because we know, independently of NRT representationalism, that the experiences we have when looking at objects that have R650 are experiences of red, but this is knowledge we have above and beyond what NRT representationalism tells us. By itself, NRT representationalism does not give the phenomenal description of the state associated with tracking R650.

A parallel problem arises if we try to generate predictions in the mind-to-matter direction. If NRT representationalism is true, then experiencing red is associated with tracking redness. But "tracking redness" is not a physical description if "redness" is understood in a way that speaks to the phenomenology (if it means what it means as part of the phenomenal description "experiencing redness"). Without being given a physicalist theory of redness, which cannot be inferred from NRT representationalism, we don't know what is the correct physical description of redness or tracking redness.

The problem can be put slightly differently as follows. A solution to the

mapping problem needs to specify a function that generates predictions such as this for all physically possible conditions that result in some conscious experience:

X experiences redness iff X stands in R to R650

Since phenomenal and physical descriptions are distinct, the representationalist schema does not have the right form to specify such a function. The following is the right form, where f is a function that maps physical descriptions of physical properties to phenomenal descriptions of experienced qualities:

An individual experiences f(Q) just in case they stand in R to Q.

Alternatively, the positions of "Q" and "f(Q)" could be swapped and the inverse of f used. Only a theory that fits this schema could possibly relate phenomenal descriptions with (distinct) physical descriptions, but I am not aware of any proposed non-trivial specification of f. So far, NRT representationalists have effectively assumed that it is the identity function.

There are other relational views of experience that might seem to specify a general psychophysical mapping of the form of the representationalist mapping, for example, naïve realism. The very form of that mapping guarantees that these theories do not supply a general psychophysical mapping.

It turns out, then, that NRT representationalism and other relational theories of consciousness do not specify a general psychophysical mapping. This is why I said earlier that a general psychophysical mapping has not even

been imagined. What we have imagined is a theory that *seems* to specify a general psychophysical mapping.

Once the apparent solution to the mapping problem offered by NRT representationalism and similar views is set aside, the problem seems completely hopeless for matter-to-mind nomic dualism. There is simply no discernible suitably general pattern in the known phenomenal-physical correlations revealed by neuroscience or everyday observation, and there is little hope of finding any.

This leaves the dualist with two choices regarding causation in the matter-to-mind direction: acausalism and anomalism. Matter-to-mind acausalism would require something like a pre-established harmony to keep mental states and physical state in sync, which is extremely implausible (especially without theism as a supporting hypothesis). It seems, then, that matter-to-mind anomalism should be considered seriously.

4 Anomalous panpsychism

In section 2, we saw that, despite obvious difficulties, the most reasonable approach to the problem of mental causation for the dualist is to reject mind-to-matter nomism in favor of mind-to-matter anomalism. The mapping problem discussed in section 3 is a problem for dualism combined with matter-to-mind nomism, and we have seen that matter-to-mind anomalism may be the best way out of this problem for the dualist. Taken together, these

considerations make a case for considering anomalous dualism, the kind of dualism that endorses anomalism about both mind-to-matter and matter-to-mind causation. As far as the considerations pertaining to mental causation and the mapping problem are concerned, anomalous dualism seems to be the most promising dualist view.

Even if anomalous dualism combines the best answers to the arguments from completeness and the mapping problem, this does not mean that the view is plausible. In particular, a number of objections need to be addressed. I will consider the following objections:

First, even if anomalous dualism is technically speaking consistent with mental efficacy, does it not give up on a truly significant causal role for consciousness? As we noted in our discussion of quantum theories of mental causation, quantum mechanics and the total body of evidence concerning the dynamics of physical systems do not seem to leave much room for mind-caused random effects to make much of a difference to the course of physical events. For this reason, it seems that mind-to-matter anomalism is inconsistent with macroscopic mind-to-matter causal connections, for example, with the fact that my conscious intention to raise my arm seems to cause my arm to raise. I will refer to the claim that there is macroscopic mind-to-matter causation but hardly any observable macroscopic random events as *causation without randomness*.

Second, anomalous dualism seems in tension with the existence of psychophysical correlations. Even if neuroscience has not solved the mapping

problem, it has uncovered numerous correlations between brain activity and conscious activity: the same brain activity is always accompanied by the same conscious activity. Call this fact *the mind-brain correlation observation*. It is not clear that anomalous dualism is consistent with the mind-brain correlation observation.

Third, if mind and matter were only randomly associated, wouldn't our experience be a mere "blooming, buzzing confusion"? Instead, our conscious minds seem to have some sort of internal cohesion. Call this fact *phenomenal cohesion*.

Lastly, anomalous dualism might not have the problem of specifying a general psychophysical mapping, but is this not simply because it gives up on explaining consciousness altogether? How could anomalous dualism explain consciousness and shed light on its place in nature without giving us psychophysical laws that explain how phenomenal states arise and what effects they have?

Despite appearances, anomalous dualism is at least in principle consistent with causation without randomness, mind-brain correlations, and phenomenal cohesion. It is also easy to see how it can be genuinely explanatory. The following picture can serve as a kind of proof of concept, though its details are obviously implausible. Suppose that physical and phenomenal properties are wholly distinct, but that every physical property is "linked" to a randomly selected phenomenal property (within certain constraints) the first time it occurs in the history of universe. As an example, suppose that the following

two constraints apply to this random linking: i) if physical properties P and Q are both instantiated at the time of the linking, they are linked to consistent phenomenal properties (assuming representationalism, we can say that two phenomenal properties are consistent when their contents are consistent); ii) if physical property P necessitates physical property Q, then P's phenomenal property necessitates Q's. Once linked, phenomenal and physical properties forever co-occur across the universe. Suppose also that the phenomenal properties of a physical system can in some circumstances have a random effect on the dynamics of the system. Say, for example, that any physical system about to enter a total physical state involving physical properties associated with inconsistent phenomenal properties randomly jumps to another physical state with a probability determined by the physical state of the system. (Such an event could be modeled as a measurement by quantum mechanics, but it doesn't have to be something that we would intuitively describe as a measurement and, conversely, events that we describe as measurements don't have to occur in this way). These suppositions together specify a view I am going to refer to as *the random theory*. This view is a kind of nonreductive anomalous panpsychism: it combines dualism (nonreductionism) with generalized anomalism and *panpsychism*, the view that phenomenal properties pervade the physical universe. I am going to refer to this kind of view more simply as *anomalous panpsychism*.

The random theory is quite implausible on its face, and seriously vague and underspecified, but it is useful to consider as a first step into the largely

unexplored conceptual space of anomalous dualism and anomalous panpsychism. I want to suggest that its principles could potentially be precisified in a way that might yield an explanation of consciousness consistent with the mind-brain correlation observation, phenomenal cohesion, causation without randomness, completeness, and the overall body of evidence regarding the dynamics of physical systems.

The mind-brain correlation observation is the easiest case: it only requires that the same phenomenal and physical properties tend to occur together. The random theory is not only consistent with this fact, but its linking principle offers an explanation for it (however intrinsically implausible it might seem as stated).

Phenomenal cohesion requires us to think about what would have happened over time if the random theory were true. Let us assume that there is some principled way of delineating the physical systems to which the random theory refers that counts properly functioning animal brains, or at least big parts of animal brains, as whole systems.[10] Under this assumption, we would expect the brains of organisms to have evolved so that they barely ever enter physical states that have been linked to inconsistent phenomenal properties

[10]Someone attracted to something like the random theory might speculate that quantum entanglement is what delineates systems. A number of authors have explored entanglement-based explanations of the fact that consciousness seems to unify contributions from different parts of a physical system (e.g. Lockwood 1989, Penrose 1994, Seager 1995). There is a widespread misconception (among philosophers) that "decoherence" virtually eliminates entanglement from the macroscopic world, but in fact the theory only predicts that decoherence makes entanglement unnoticeable by making the results of quantum measurement statistically like those of classical measurements. See Schlosshauer 2005 for a relatively non-technical explanation of decoherence.

(or at least not macroscopic states linked to inconsistent properties). Unless they had so evolved, they would be suffering from the disruptive effects of random jumps, which would make them unstable, hence not prone to survive. Regarding simpler physical systems (including, in the limit case, isolated particles), their having fewer phenomenal properties (in virtue of having fewer physical properties) might explain their stability: they are not very likely to enter inconsistent states. The final result, then, should be brains and other macroscopic physical systems in which the potential random effects due to consciousness are largely absent, which is what we find. We should also expect the resulting stream of experiences supported by human brains to be generally coherent, which is what we find (phenomenal cohesion).

In addition to predicting a certain orderliness that is consistent with what we find, the random theory also predicts that phenomenal-physical associations should appear essentially random except for any structure implied by the linking principle. Again, this is roughly consistent with what we find. This is particularly interesting because no other theory even begins to explain the apparent randomness of mind-body correlations. Our half-baked anomalous panpsychism seems to provide not only a proof of concept for an explanation of the ways in which consciousness is organized, but it seems that a story along these lines could also potentially explain the ways in which it is *dis*organized. This might answer the charge that anomalous dualism gives up on explanation.

We have yet to show that anomalous dualism, anomalous panpsychism, or

the random theory is consistent with causation without randomness, which one might say is the main problem with these views.

There is one important way in which the random theory illustrates the possibility of macroscopic effects consistent with completeness and a general lack of macroscopic quantum effects. If this theory is true, consciousness might have played a role in structuring physical systems in the past, progressively weeding out complex systems whose tendency to generate inconsistent experiences makes them unstable. This requires that there were many macroscopic random events triggered by consciousness in the past, but not that such events be common today. In this way, the random theory illustrates the possibility that consciousness played an important causal role in the course of evolution without making a detectable difference today.

This addresses the need for some efficacy throughout natural evolution, but this does not fully address the objection that consciousness has macroscopic effects that are not simply random effects. In particular, this does not accommodate alleged causal connections such as a mental state causing a bodily movement.

The first thing to note here is that it is almost certain that no mental state is nomologically sufficient for a bodily movement. The most that we can ask for is causal *relevance*, not strictly speaking *causation*.

Without giving an analysis of causal relevance, it is plausible that events that stand in counterfactual dependence relationships are *in some sense* causally relevant to each other. For example, had there not been a spark, the

fire would not have started. This seems to make the presence of the spark causally relevant to the fire.

A view along the lines of the random theory could potentially deliver just this kind of causal relevance between mental events and bodily movements in a manner that is entirely consistent with completeness and the general absence of macroscopic random events. To illustrate, take the following counterfactual claim:

Chocolate Had you not consciously thought that there was a chocolate bar in front of you, you would not have reached out in the way you did at the time you did.

If Chocolate were true, your conscious thought that there was a chocolate bar in front of you would be causally relevant to your reaching in the way you did, just like the spark is causally relevant to the fire. The random theory can potentially be precisified in such a way as to be consistent with, and explain, facts such as Chocolate consistently with all evidence on hand. Suppose, for example, that your perceptual systems were rigged in such a way that, given their input at t (a retinal image of a chocolate bar), they will trigger a conscious thought to the effect that there is a chocolate bar in front of you at a specific location L if you are not having such a thought at t. At the same time, suppose that your cognitive system was poised to theorize about what is in front of you in such a way that if you don't consciously think that there is a chocolate bar in front of you at t, you will quickly form a conscious thought to the effect that there is a certain non-chocolate desert,

say, lemon pie, in front of you at L. In sum, your brain is in such a state that, as a matter of nomologically necessity, if you are not having a thought that there is a chocolate bar at L at t, it will, at t+1, produce both a conscious thought that there is a chocolate bar at L *and* a conscious thought that there is lemon pie at L. Given that these thoughts have inconsistent contents, this setup makes the following counterfactual true: had you not consciously thought that there was a chocolate bar in front of you, you would have formed inconsistent conscious thoughts. On the random theory, inconsistent phenomenal states trigger random disruptions. In principle, we could flesh out the details of the case and the random theory in such a way that this would have prevented you from reaching at just the time you did. This would make Chocolate true. Of course, this example is highly contrived (I am only trying to make a point of principle), but, for all we know, it could be that the brain's massively redundant architecture ensures that many phenomenal states cannot be altered without generating an inconsistency, which would underpin a kind of causal relevance for these states on the random theory.

This kind of causal relevance could be pervasive consistently with completeness and a general absence of detectable macroscopic random events. This is because counterfactuals make no observable difference. For example, the truth of Chocolate makes no observable difference to the course of physical events. If the physical universe and the organisms it contains had evolved to keep consciousness-caused random events to a minimum as the random theory seems to predict, we would expect numerous counterfactuals

like Chocolate to be true, which would give consciousness widespread causal relevance that makes no detectable difference today.

While anomalous dualism is consistent with two important causal roles for mental events, there are also causal roles that it is not consistent with: it does not allow mental events to be nomologically necessary nor sufficient for physical events, and, intuitively, it does not allow causal *oomph* to pass from mental events to physical events in a deterministic way. One might say that these are important shortcomings of the view. They are perhaps shortcomings, but anomalous dualism at least succeeds in accommodating the core evidence regarding mental causation. We don't have very strong reasons to think that causal *oomph* passes between mental and physical events, much less to think that the two kinds of event are subsumed under nomological principles. Such claims are simply not apparent to ordinary or scientific observation. The key reason mental causation must be accommodated is not that it is directly observable, but that it seems necessary in order to make sense of the place of consciousness in the mind and nature, in particular, of the fact that brains seem to have evolved to make use of conscious states in some way (as noted in section 2). Anomalous dualism promises to satisfy on this score.

Our exploration of the random theory suggests that a view along these lines can in principle explain the apparent arbitrariness of phenomenal-physical associations while being consistent with macroscopic mental efficacy, an evolutionary role for consciousness, the existence of numerous mind-brain cor-

relations, phenomenal cohesion, the completeness of physics, and a general lack of observed macroscopic randomness. As a kind of dualism, the random theory is also consistent with arguments against all types of physicalism (including those of Chalmers 1996 and Goff 2009). This makes anomalous panpsychism the only position on the mind-body problem that is not currently open to principled objections (that I know of). Of course, the random theory is implausibly vague and almost certainly not entirely correct. It is an open empirical question whether or not a more precise anomalous panpsychic theory can be specified that retains the theoretical virtues of the random theory and is consistent with everything we know or may find out in physics and neuroscience. My goal in this paper was only to try to open some new conceptual space. Even though the random theory itself is implausible, its many virtues suggest that the overall approach of anomalous panpsychism might deserve more investigation.

References

Bennett, K. (2007). Mental causation. *Philosophy Compass*, 2(2):316–337.

Bourget, D. (2004). Quantum leaps in philosophy of mind. *Journal of Consciousness Studies*, 11(12):17–42.

Bourget, D. (2010). Consciousness is underived intentionality. *Noûs*, 44(1):32–58.

Bourget, D. (2015). Representationalism, perceptual distortion and the limits of phenomenal concepts. *Canadian Journal of Philosophy*, 45(1):16–36.

Bourget, D. (2017). Representationalism and sensory modalities: an argument for intermodal representationalism. *American Philosophical Quarterly*, 53:251–267.

Bourget, D. (forthcoming). Why are some experience "vivid" and others "faint"? representationalism, imagery, and cognitive phenomenology. *Australasian Journal of Philosophy*.

Bourget, D. and Mendelovici, A. (2014). Tracking representationalism. In Bailey, A., editor, *Philosophy of Mind: The Key Thinkers*, pages 209–235. Continuum.

Byrne, A. (2001). Intentionalism defended. *Philosophical Review*, 110(2):199–240.

Chalmers, D. (2015). Panpsychism and panprotopsychism. In Alter, T. and Nagasawa, Y., editors, *Consciousness in the Physical World: Essays on Russellian Monism*. Oxford University Press.

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

Chalmers, D. J. (2004). The representational character of experience. In Leiter, B., editor, *The Future for Philosophy*, pages 153–181. Oxford University Press.

Churchland, P. S. (1986). *Neurophilosophy: Toward A Unified Science of the Mind-Brain*. MIT Press.

Crane, T. (2003). The intentional structure of consciousness. In Smith, Q. and Jokic, A., editors, *Consciousness: New Philosophical Perspectives*, pages 33–56. Oxford University Press.

Davidson, D. (1970). Mental events. In *Essays on Actions and Events*, pages 207–224. Clarendon Press.

Eccles, J. C. (1978). *Mind and Brain*. Paragon House.

Goff, P. (2009). Why panpsychism doesn't help us explain consciousness. *Dialectica*, 63(3):289–311.

Lewis, D. (1979). Counterfactual dependence and time's arrow. *Noûs*, 13(4):455–476.

Lewis, D. K. (1973). *Counterfactuals*. Blackwell.

Lockwood, M. (1989). *Mind, Brain, and the Quantum*. Oxford University Press.

Mendelovici, A. (forthcominga). Panpsychisms' combination problem is a problem for everyone. In Seager, W., editor, *The Routledge Handbook to Panpsychism*.

Mendelovici, A. (forthcomingb). *The Phenomenal Basis of Intentionality*. Oxford University Press.

Papineau, D. (2001). The rise of physicalism. In Gillett, C. and Loewer, B. M., editors, *Physicalism and its Discontents*. Cambridge University Press.

Pautz, A. (2009). A simple view of consciousness. In Koons, R. C. and Bealer, G., editors, *The Waning of Materialism: New Essays*, pages 25–66. Oxford University Press.

Penrose, R. (1994). *Shadows of the Mind*. Oxford University Press.

Schlosshauer, M. (2005). Decoherence, the measurement problem, and interpretations of quantum mechanics. *Rev. Mod. Phys.*, 76:1267–1305.

Seager, W. E. (1981). The anomalousness of the mental. *Southern Journal of Philosophy*, 19(3):389–401.

Seager, W. E. (1995). Consciousness, information, and panpsychism. *Journal of Consciousness Studies*, 2(3):272–88.

Seager, W. E. (1999). *Theories of Consciousness: An Introduction and Assessment*. Routledge.

Speaks, J. (2010). Attention and intentionalism. *Philosophical Quarterly*, 60(239):325–342.

Speaks, J. (2015). *The Phenomenal and the Representational*. Oxford University Press.

Stapp, H. P. (1996). The hard problem: A quantum approach. *Journal of Consciousness Studies*, 3(3):194–210.

Yablo, S. (1992). Mental causation. *Philosophical Review*, 101(2):245–280.