Materialism, Then and Now

Peter Godfrey-Smith
University of Sydney

Written for a conference at the University of Sydney to mark the 50th anniversary of David Armstrong's *A Materialist Theory of the Mind* (1968), August 2018. The final version will appear in a collection of papers from the conference edited by Peter Anstey and David Braddon-Mitchell, called *A Materialist Theory of the Mind: 50 Years On*, with Oxford University Press.

1. Introduction

David Armstrong was the Challis Professor of Philosophy when I arrived as a first year undergraduate at the University of Sydney in 1983.[1] He taught epistemology in first term. (Bill Lycan, visiting from the USA, handled metaphysics). I was told that Armstrong taught first year regularly because he thought that all students should be taught at some stage by "The Professor." The lectures were good – patient, considered, and fair. They also encouraged us to see philosophical debates in rather competitive, combative terms ("spoils to the victor," "a problem for everyone is a problem for no one" – I am not sure if I heard those phrases in that first class, but certainly early on). In my fourth year, Armstrong taught a whole-year course that was essentially the writing of *A Combinatorial Theory of Possibility* (1989). I was not entirely on board with that project

[1] The text here is fairly close to the talk given at the Armstrong conference in August 2018, but I have filled out the last part of each of the main sections, and incorporated a couple of points made in discussion. Other points have been amplified in footnotes. Thanks to Peter Anstey and David Braddon-Mitchell for organizing the conference, and for helpful comments on an earlier draft.

and its methods, but enjoyed the course a great deal and relished, as always in the T&M department in those days, the sense that I was seeing philosophy that mattered being developed and shaped in real time.

The T&M or "Traditional and Modern" Philosophy Department in which Armstrong taught was one of two at Sydney, born of a split in the 1970s. During my time at Sydney in the 1980s, a lot of work in the philosophy of mind was being done within a more American style, one imported in part by Michael Devitt, who had returned from Harvard some time before, a style fostered also by American visitors like Bill Lycan and Stephen Stich. We talked about mental representation, intentionality, and computation, reading Jerry Fodor and Dan Dennett (and in my case, Fred Dretske and Ruth Millikan). Armstrong seemed to some of us a bit stodgy in this context, and he did push back against that. His annoyance was sometimes evident, especially at the idea that there was a history that ran from "the identity theory" to "functionalism," with him as the last part of the first, superseded stage. He denounced "trendiness" in naturalistic philosophy. He was, again, a combative sort of person. This affinity with conflict served him in good stead during the creation of the T&M department, which was, as David Stove always said, a wonderful place to be. It enabled Armstrong to get through a period in the 1970s when he was actively shunned and excluded from social events – with groups leaving lunch tables at the staff club when he sat down.[2] Some of those people later benefited greatly from the birth of the T&M Department, and I hope they apologized.

Armstrong was in other ways modest. I heard him say more than once that his hope was to be footnote to the history of philosophy. I think it is becoming clear that it is only in the briefer histories that he is just a footnote. He also endorsed the "under-laborer" description of philosophy, which was surprising to me. Perhaps he thought that if this was good enough for Locke, it was good enough for him. But this is obviously inaccurate, as is the official emphasis on conceptual analysis in *A Materialist Theory of the Mind*. The book is full of material that is more substantive, more abstract-empirical, as he later admitted in the preface to the paperback edition.

[2] I learned this recently from Kim Sterelny.

My first topic in this chapter is how Armstrong set up his materialism, especially in *A Materialist Theory of the Mind* ("*MTM*," 1968). Then I'll say something about how we should think about materialism now.

2. Materialism Then

Armstrong's view is expressed as a "causal theory," a causal analysis of mental states and processes. It is a view that grants quite a lot to behaviorism, while holding that mental states are inner states that act as causes. We get a handle on mental states via their causal role, especially as causes of behavior. Science can then tell us that particular neural states and processes are the things that do that causing, the things that play these roles in the actual world. As a result, we can identify mental states and processes with those neural states and processes.

> The concept of a mental state is primarily the concept of *a state of the person apt for bringing about a certain sort of behaviour*. Sacrificing all accuracy for brevity we can say that, although the mind is not behaviour, it is the *cause* of behaviour. In the case of some mental states only they are also *states of the person apt for being brought about by a certain sort of stimulus*. But this latter formula is a secondary one. (1968, p. 82)

Two things he does not worry much about are the modal status (necessity/contingency) of his claims, and multiple realizability. In the paperback preface he says he accepts multiple realizability and always did, once the question was raised.

In retrospect, I think we can see Armstrong's project as follows. He wants to cover the whole of the mind (near enough), but he does this by working his way in from two particular parts of the mind that get a great deal of attention. These are *perception* (and its relatives), and states related to the *will* (such as intentions). Both these aspects of the mind are *peripheral,* in a sense. When I say that, I don't mean they are secondary in importance, but peripheral in a more literal sense. They are close to the mind's interaction with the non-mental. Armstrong does not present them this way, or give this as the reason for his focus on them. He just says they are important. He says more than once that

perception and the will are "great" and "central" topics for the philosophy of mind.[3] He sees himself as going after what is basic. But we can also see their peripheral character, on each side. These peripheral states can be contrasted with what I will call the *core*. The core is belief-desire interaction – the way that actions stem from what you want in the light of what you believe – together with the role of memory and other faculties that bear on belief-desire interaction. Perception feeds beliefs, and the belief-desire nexus gives rise to intentions and actions.

The shape of belief-desire interaction, and how to model it rigorously, were seen clearly by Ramsey in "Truth and Probability" (1926) – perhaps by others even earlier, but certainly by Ramsey. I will say more about Ramsey's achievements in relation to Armstrong below.

In the light of Armstrong's emphasis on perception and intention, we can see him as situated between and responding to British empiricism on one side, and on the other side, Rylean behaviorism and Wittgensteinianism. The empiricist influence is clear: Locke and Hume are often discussed, and sometimes also Berkeley. But Armstrong was also impressed with Ryle, and the achievements of the people around him and Wittgenstein. *MTM* includes lavish praise for Wittgenstein and behaviorism – talk of "astonishing progress" and "epoch-making philosophical achievements." Whatever the final view of the mind is, Armstrong says, it must make peace with behaviorism.[4]

He took seriously both of these very different views: *Empiricism*, with its individualist and first-person character, with its proneness to skepticism and phenomenalism, and its running into trouble with more cognitive aspects of the mind; and *behaviorism*, with its social, third-person viewpoint, and its deflationary outlook on the tradition of sensationalist empiricism. Armstrong was influenced by both, and each has their favorite mental states: perception on the empiricist side, and the states most closely related to action, such as intention, on the other side.

A consequence of this approach, though, is a failure of a causal theory of the mind of Armstrong's kind. The core, the belief-desire nexus, resists an analysis of the sort he gives, and resists it in part *because* it is non-peripheral. I'll discuss this point using the

[3] See *MTM* p. 187 and p. 245.
[4] See *MTM* p. 67 and p. 118.

case of belief, first through a direct application of Armstrong's summary above, and then with a look at what Armstrong explicitly says about belief in a late chapter.

Suppose we apply Armstrong's formula directly. If you have a particular belief, you are disposed to do certain things; there are behaviors the belief is apt to cause. This would be a modification of the Rylean idea that to believe something is to have (complex) dispositions to behave in certain ways. The problem is that the effect on action of any belief depends on other beliefs and on desires. Two people who both believe that Baltimore is in Maryland could behave indefinitely differently, depending on what they want and what else they believe. Even if they are asked a yes-or-no question in a quiz show with $1 million at stake, they will answer differently if one thinks they are being tricked, or that the money will corrupt them, or whatever. The thing that causes an action is a complex of beliefs and desires.

Armstrong does not want to be simplistic about the causal role of core states. He looks for ways to hedge and qualify what is claimed. That would be OK if there was a central tendency in causing behavior that each belief had, one that might get interfered with in various ways. But there is no central tendency in a belief's role. Core states are just not amenable to a treatment of this kind. The idea of "an inner state bringing about outward behavior" (1968, p. 80, emphasis added) just does not apply to beliefs.

There is a standard objection to Ryle and logical behaviorism called the "Chisholm-Geach" objection, which holds that attempts to describe the dispositions associated with a particular mental state will collapse into circularity, because the behaviors associated with the belief that *p* will depend on other mental states. As Lycan noted in discussion, my point here is that the Chisholm-Geach objection is as much an objection to Armstrong as to behaviorism.[5] Armstrong's view is close enough to behaviorism to inherit one of its main problems.

---

[5] See Chisholm 1951, and Geach 1957. Here is Geach: "[I]s there in fact any behaviour characteristic of a given belief? Can action be described as "acting as if you held such-and-such a belief" unless we take for granted, or are somehow specially informed about, the needs and wants of the agent? In Ryle's example this information is smuggled in by his speaking of a *gardener's* rain-expecting behaviour (and tacitly assuming that the gardener is not e.g. a discontented or corrupt servant who wants the garden to be ruined). When Dr Johnson did penance in Uttoxeter market-place, he may have begun by standing around bareheaded until the threatened shower should fall; this would not be recognizable as rain-expecting behaviour without a knowledge of Johnson's wish to do penance." (1957, p. 8)

In understanding core states, there is no way of getting by with less structure than a belief-desire profile, of the sort Ramsey worked with in "Truth and Probability" (1926). Here is Ramsey:

> I propose to take as a basis a general psychological theory, which is now universally discarded, but nevertheless comes, I think, fairly close to the truth in the sort of cases with which we are most concerned. I mean the theory that we act in the way we think most likely to realize the objects of our desires, so that a person's actions are completely determined by his desires and opinions. This theory cannot be made adequate to all the facts, but it seems to me a useful approximation to the truth particularly in the case of our self-conscious or professional life, and it is presupposed in a great deal of our thought. It is a simple theory and one which many psychologists would obviously like to preserve by introducing unconscious desires and unconscious opinions in order to bring it more into harmony with the facts. (p. 173)

That move is not yet pushing us far from an Armstrong-type view. What does push away from it is Ramsey's use of an expected utility model, again as an approximation: you will do what is associated with the best expected payoff, all things considered. That requires that degrees of belief be recognized; this is part of what makes it possible to use very small differences in behavior (such as whether you would walk 100 yards or 120 yards out of your way to ask directions) in getting a handle on mental states. Behavior stems from degrees of belief and from evaluations of events that are the possible outcomes of actions.[6] (Again, Ramsey keeps saying this is an approximation, but the ways in which it is an approximation do not take us back towards an Armstrong-like view.) If this model is accepted, what remains of belief as "an inner state bringing about outward behavior"? A (degree of) belief might be a function (in the mathematical sense) from contexts to behavioral dispositions, where each "context" is much of the rest of a psychological profile. That is the closest a belief can get to being disposed to issue in some particular behavior.

---

[6] Ramsey: "I suggest that we introduce as a law of psychology that his behaviour is governed by what is called the mathematical expectation; that is to say that, if *p* is a proposition about which he is doubtful, any goods or bads for whose realization *p* is in his view a necessary and sufficient condition enter into his calculations multiplied by the same fraction, which is called the 'degree of his belief in *p*'. We thus define degree of belief in a way which presupposes the use of the mathematical expectation" (p. 173).

What about desires? Do they have the same "core" properties? Is it true that a desire can issue in almost any behavior, depending on what is believed and what else is wanted? Armstrong has a long discussion of forms of desire that are removed from causing acts directly, but action-causing remains central: "Desires are essentially action-producing although they may be inhibited or prevented from producing action by other desires" (p. 151). Are there some desires that are "apt for causing" actions in a straightforward way? Or are the effects of all desires mediated by beliefs, with the same holistic features I associate with the core and the Ramsey model?

I think desires are a bit special. When desires are directed at outcomes distinct from acts, then their role is always indirect and belief-mediated. I desire to get to work on time. That could result in almost any act, depending on my beliefs. But a few desires seem to be aimed at acts alone. I desire to shut my eyes (not: that my eyes end up closed). This desire does not apparently have the holistic features. "Desire" then covers both core and non-core states.[7]

I said above that the core was described by Ramsey, decades before Ryle and Armstrong. Ramsey was remarkably philosophically independent. You might say that he was not around along enough for the traditions to seep into him, but he did go quite far down his own distinctive path. To a significant extent, Ramsey made the core unmysterious, by showing how subtle differences between belief-desire profiles can be made manifest. I have to say all this in a qualified way; Ramsey did not give a behaviorist reduction of the core. The bridging components in his analysis are dispositions to choose between big hypothetical scenarios, and between gambles over those scenarios. (He says things like this: "If then we had the power of the Almighty, and could persuade our subject of our power, we could, by offering him options, discover how he placed in order of merit all possible courses of the world" p. 176.) As Ed Elliott (2015) discusses, this does not really forge a constitutive link to behavior. The same holistic factors discussed above arise for behavioral expressions of these preferences, too. The subject's mental "placing in order" and their behavioral "placing in order" need not simply align.) We might understand the Ramsey project as using an idealization, making use of a peripheral

[7] This might be called an "intention" rather than a desire. I think all the language around here is pretty flexible.

mental state that has special properties (reminiscent of Armstrong), or the project might be carried forward differently. But there is a lot of progress made in handling the behavioral manifestation of core states.[8]

Armstrong – like nearly everyone, with the partial exception of Lewis – was not looking to Ramsey *in this way.* Not even Ryle was.[9] It is odd how little connection there was, for a long time, between Ramsey's discussion and the rest of philosophy. Lewis in "Psychophysical and Theoretical Identifications" (1972) gave an account that can apply to the core. That account used the idea of a "Ramsey sentence," which is derived from a different Ramsey paper.[10] In his 1972 paper, Lewis did not discuss *how* belief-desire profiles might be made manifest, and did not note what Ramsey had done in "Truth and Probability." He just discussed the general way that Ramsey sentences might be used to make sense of a complex network of causal roles and the occupants of those roles.

I now return to Armstrong. Armstrong discusses belief at various points in the central chapters of *MTM*, but as he is focused so much on perceptions and intentions in those chapters, belief is not handled in detail. Then in a short late chapter, he addresses belief specifically. Here he says that *perceptual beliefs* are basic, and those are understood in terms of their enabling discriminative behaviors, and also says that perceptual beliefs (taken together) are like a big "map" of the environment. *Non*-perceptual beliefs (beliefs held today about Caesar, and so on) are *extensions* of the map. Those beliefs are a "a 'map' of reality that is an extension of the 'map' formed by our perceptual beliefs" (p. 343). What sort of extension? As he says both in 1968 and the later paperback preface, his discussion here is very thin. Suppose one took it seriously and as

---

[8] Elliott (2015) looks at where and why various other attempts fall down on the most ambitious task here, showing how pure behavioral dispositions might reveal preferences and degrees of belief. Elliott's treatment has glass-half-empty character, which should not be seen as the whole story, though I concede that I don't know how to get more into the glass.

[9] Ramsey is not in Geach's 1957 book either. There's a question about the history of the field that would be worth a close look here. There seem to be some sort of subtle but powerful framing effect at work, where Ramsey's discussion is often not perceived as relevant. I was not thinking about it in the present connection until I looked at the connections between Peirce and Ramsey, and then encountered Elliott's work.

[10] Apparently the importance of this idea was recognized first by Hempel. Here I draw on Psillos (2004), who says that Hempel coined Ramsey-sentence' in his 1958 "The Theoretician's Dilemma," drawing on Ramsey's posthumous 1929 "Theories." Psillos says that Braithwaite had used Ramsey's paper a bit, and Carnap had reinvented the idea before he read Hempel's paper.

literally as seems possible. Then, insofar as our beliefs are map-like, they would affect action in something like the way seen in Ramsey's treatment: actions are determined by the big map (with confidence levels) together with evaluations of the likely effects of actions. We would end up far away from the idea that beliefs are apt for causing particular behaviors. The "map" idea is taking us down a different path.[11]

Armstrong at other places is explicitly tries to qualify his "apt for causing" formula in the case of things like beliefs. A perceptual belief does not really issue in particular behaviors, he says, but *makes possible* some particular behaviors, which are discriminative behaviors. You might not want to perform those behaviors, but you can if you want.

> [I]n the case of perception, there is no question of the inner event actually tending to bring about behavior. What we must say, rather, is that perception supplies a necessary precondition for appropriate behavior. If the baby can perceive the difference between a green and a blue block, then it is in a position to discriminate between them in its behavior if it should want to.
>
> Perception can then be compared to acquiring a key to the door. If I acquire such a key, I have acquired an instrument which may be a necessary precondition for the opening of the door, even if I subsequently ignore the door completely. (p. **)

However, to say this is to completely abandon the apt-for-causing analysis, the approach guided by ideas like this:

> [I]f we say that the baby's perception is the coming to be of a state of the baby apt for the bringing about of certain sorts of discriminative behavior, we get the best of both realms. We preserve the 'inner' character of perception, yet at the same time we create a logical tie between the inner event and the outward behavior. (p. 248)

Armstrong may have thought that if we drop that last idea, we are doomed; we have to hang onto some version of that last in order to give any causal analysis of the mind at all.

[11] This path is explored in some detail by Braddon-Mitchell and Jackson (2006).

But Ramsey had made progress on this problem decades before. He showed that if you do the *double* complication – from simple belief to degrees of belief, and from individual states to whole profiles – you can go a long way towards making sense of the behavioral manifestation of core psychological states. Those complications make things better, not worse.[12]

Before moving on, I'll say something about Armstrong's treatment of belief in the most relevant work after *MTM*, which is *Belief, Truth, and Knowledge* (1973). This book expands on the "map" analogy for beliefs, and this involves yet another distinct link to Ramsey. Ramsey in "General Propositions and Causality" (1931) talked of belief as comprising a "map... by which we steer." In the 1973 book Armstrong takes this on board, and accepts, as I read him, that this moves us away from the kind of causal analysis given in *MTM*.[13]

On this new road, different questions arise. Map-talk might be understood here as referring to the belief side of the belief-desire nexus as handled in an expected utility model. That leaves open the possibility of a holistic determination of the belief-desire profile of an agent at a time. Alternatively, the "map" idea could point towards inner states with a particular kind of combination of representational and causal properties – picture-like representations that we use to help us steer. Perhaps there is a place for both. We are thus drawn into the kinds of debates about intentional psychology and mental representation seen in the 1980s. Are there "explicit" mental representations, and how do they work? The idealizations made in an expected utility model now start to matter. Which features of ordinary belief-desire interpretations are aimed at picking out inner causes, and which ones have a "looser fit" to what is going on inside?[14]

---

[12] Another way to put it: without degrees of belief, the "Chisholm-Geach" phenomenon is just a problem, something that messes up various reasonable-looking projects. With degrees of belief, which Armstrong and others set aside as a secondary matter, the core becomes something that can be represented and studied. What makes things look harder ultimately makes them easier.

[13] Armstrong: "It is an essential mark of beliefs, then, as opposed to mere thoughts, that, if suitable dominant desires are also present, the believer is moved to action: action having as its objective the satisfaction of the desire. This does *not* mean that an account can be given of beliefs as being nothing but causal factors in the believer which initiate and sustain certain courses of action in certain circumstances." (1973, p. 72).

[14] This last idea is followed up in detail by Dennett in his neo-Rylean works (1971, 1987).

3. Materialism Now

The second half of the chapter will look at different themes: at what materialism *is*, and why we should believe it.

What is it to be a materialist? Armstrong hardly discusses this question in *MTM*, taking it for granted that if you identify mental states with "physico-chemical states of the nervous system" (p. 356), then you are a materialist, at least about the mind. But recent discussion has shown that none of this is straightforward. The *material* or the *physical* is not a very clear category. (I will often use the term "physicalism" from here, but I don't intend this as a significant shift.[15]) We shouldn't commit to materialism or physicalism by saying that textbook physics in its current form gives us an inventory of everything that exists, or the low-level basis of everything that exists. Present-day physics is probably not entirely accurate. It can be tempting to say that to be a physicalist is to think that a future and completed physics will describe everything that exists (or its low-level basis), but that can't be the right approach, either. If it turns out that God, or a World Spirit, exists, the Physics Department can be expected to take an interest. A future physics might not be physicalist, in the sense on the table here. Perhaps a physicalist is committed to the idea that the physics of the future will be a clear descendant of current views, which are on the right track? This is what Lewis (1994) said. I think this is wrong, too. We can be confident, perhaps, that future physics will agree with present physics about a large range of macroscopic phenomena, but are we so sure it will agree on what sorts of things are fundamental? Lewis said:

> We may think, for instance, that mass and charge are among the fundamental properties; and that whatever fundamental properties remain as yet undiscovered are likewise instantiated by very small things that come in very large classes of exact duplicates. (1994, pp. 51-52).

This is too strong. Suppose the physics of the future uses a field ontology rather than a particle ontology, and treats particle-like phenomena as derivative. That is not very

[15] Lewis (1994) says that it would be pedantry, or worse, to switch from "materialism" to "physicalism." Both terms do have problems. A modern version of "materialism" is as much about energy as matter, as Lewis notes, and "physicalism" suggests a relationship to a field of study rather than a kind of entity, something I think would be a mistake.

farfetched – various physicists have believed it. It must be possible for a future field-based view to count as physicalist.

How do we steer through this? Is there a view of materialism or physicalism that is broad enough to allow surprises that involve big holistic fields, but narrow enough to rule out a future physics of spirit? We should accept that the category of "the physical" has become a bit unclear, but I think there is a steering-through that is possible. The heart of my own view is: (i) monism, and (ii) the construction of mind from things that are not mental (and have no mental properties). This leaves some flexibility about what counts as a physical fact, but does not fall into panpsychism. This view is related to what is now sometimes called *via negativa* physicalism. I'll discuss towards the end whether this is strong enough to mark out materialism from all other options; I think some of the boundaries are a bit vague.[16]

Why should one believe a view of this general form? I am going to take it as a background assumption that standard anti-materialist arguments have adequate replies. Here I have in mind the "knowledge" argument, zombie arguments, and so on. Some of the standard replies are adequate in each case.[17] My main topic in this section is the positive arguments that can be given for materialist views.

A good way to make an argument of this kind is to make use of continuities between us and other organisms, with respect to the relations between physical and mental properties, and use these continuities as a constraint on overall views. I don't make the argument using parsimony. Parsimony considerations are important here for many philosophers, including Smart (1963) and Armstrong, but I regard parsimony arguments, in their (slender) reasonable forms, as inapplicable in this context. Instead, we should ask: what are the *rival overall pictures*, and which ones seem likely to be filled out in a way that works, especially when we have our eye on all sorts of living things? Dualism is most plausible when we can set things up in a way that contrasts humans with everything else. Descartes was able to use a framing of this kind. It is much harder to be a dualist when you accept psychological continuities extending through many kinds of organisms and an evolutionary story that links all these features. It is hard to be a dualist

[16] See Montero and Papineau (2005) for a discussion of *via negativa* options.

[17] For knowledge arguments, see Jackson (1982), and for zombie arguments, see Chalmers 1996). Godfrey-Smith (2019) discusses in more detail what I see as the right replies in each case.

when you want to make sense of continuities between human mental phenomena, non-human animal mentality, and the proto-cognitive capacities seen in all cellular life.

I will develop that argument further in a moment. First I'll say something about the other approach, which uses parsimony.

Parsimony arguments are prominent in Armstrong's book. They are invoked often in the final chapter, where Armstrong wraps up his case for materialism, especially when he has to deal with non-classical dualist views. For many, the problem for traditional dualist views is the problem of interaction, but weaker dualisms are harder to rule out, including epiphenomenalism and emergentism of some kinds. Parsimony deals with those.

The argument from parsimony is offered, by Armstrong and others, in a setting in which it is unclear whether there is *any* reason that simplicity is a guide to truth, on this issue or others. No positive reason is offered to believe that simplicity is a guide to truth. This is surely a major problem. What lies behind many uses of parsimony of this kind is perhaps the implicit message that if this is not reasonable, then all hell breaks loose – not just here, but everywhere. Abandoning parsimony leads to a complete falling-apart of rational theory choice, as there will always be many options compatible with our evidence. Parsimony about the mind-body problem, then, is an application of a principle that might be hard to justify, but that we can't possibly do without.

A general rethinking of simplicity preferences is needed, both in science and philosophy. Simplicity doesn't mean nothing, but it's not a guide to truth, especially in cases like this. A brief sketch of the view I'd defend is as follows. Simplicity does not give us reason to believe a theory is true, but simplicity preferences do have a role in science. First, a preference for simplicity is part of a good policy or strategy aimed at scientific change. A good rule is to start simple and expect to get pushed elsewhere. This is a broadly Popperian attitude. Suppose instead we began with a more complex theory. It is no less likely to be true than the simple one, but the process of being pushed from old to new views by incoming data is less straightforward. Simple theories are good places from which to initiate the dynamic process characteristic of theory development in science. Occasionally a simple theory might actually be true; that is merely a bonus.

Second, there can be domain-specific empirical reasons to prefer simplicity in some cases (Sober 1990). If you think that beneficial mutations are rare events in evolution, then a story that requires fewer of them has an advantage. Third, in a Bayesian context, the Popperian point above has a positive flipside, one sometimes now known as "Bayesian Occam's razor" or the "automatic Occam's razor." When a theory constrains the data a lot (takes risks, is easily falsified), if incoming data *does* fit the theory, the theory gets a lot of credit, more than it would if it had only weakly constrained the data. In some cases, simple theories do constrain the data more than related theories that are more complex. In other cases, they don't – a simple theory might also be vague about some of what is going on. Simplicity in this Bayesian context does not operate as an extra principle that gives one theory an edge over others after the data has done its job. Instead the razor, when real, is a reflection of the fact that theories that constrain the data a lot get credit when their predictions are born out, and in some cases, simple theories constrain the data a lot.[18]

None of these three scientific roles for simplicity applies to the mind-body problem.

If you want more than this from parsimony, then tell me how we can have more. Don't say: if there is no more than this, then we're doomed. We're not doomed in ordinary scientific contexts, where we are continually getting new data, and not doomed either in the philosophical cases. In that context, our project should be to get a sense of the range of genuine options, and choose the best one according to how all these options handle various kinds of evidence. The problem with many alternatives to materialism is not that they are unparsimonious, but they don't have resources with which to tell a coherent story. When they are fleshed out, they do badly with what we know.[19]

---

[18] Popper (1959) discusses the link between simplicity and falsifiability. For the "Bayesian Occam's razor," see Jefferys and Berger (1992). Some of this literature takes it as constitutive of simplicity that simple theories constrain the data more than complex ones, thereby tightening the link between simplicity and evidence, but that move is quite artificial.

[19] Lycan (2009) also criticizes appeals to parsimony by materialists. I am more skeptical about these arguments than he is. For Lycan, simplicity is a genuine tie-breaker in these sorts of debates when other things are genuinely equal (*ceteris paribus*), but genuine ties are rare and this is not one of them. I agree that ties are rare, but don't think simplicity is a tie-breaker of the kind traditionally envisaged even in a genuine tie.

Here is a more detailed outline of an argument of that kind, using the continuities between human and nonhuman organisms. The starting point is our own case – our own brains, the cognitive side of the mind, and subjective experience. There does seem to be an initial difficulty in explaining experience (consciousness, in a minimal sense) using those resources, so dualism becomes a possibility. Then we look at other animals, finding not just brains and the control of behavior, but many apparent marks of experience. We have an outline of an evolutionary story about how these capacities arise, including ones that seem to have some relation to felt experience. We can trace the evolution of sensing and the perception of objects, of the internal valuation of events, of the ongoing tracking of the divide between self and other, which is needed by a mobile animal to make sense of what it perceives (Merker 2005). All these capacities have simple beginnings – some in unicellular life, some in invertebrate animals – and there is no reason to believe in an evolutionary leap in the evolution of subjectivity.

Given all this, what are the options? If we want to opt for interactionist dualism, would we then hold that little scraps of the extra non-physical stuff enter the story at many successive stages? Or that there might be more or less limited neural communion with a domain of nonphysical mind? Alternatively, we might argue that a full-fledged version of experienced mentality comes in at some specific stage. Why would this happen? It would help if there were some *sui generis* features of human brains, or perhaps mammal brains. It is not impossible to argue for such a thing, and many sorts of empirical work might be relevant. Recent work, though, has gone the other way, and has also tended to find apparent marks of some sort of experience in an ever-wider range of animals, however (crabs, bees, octopuses...). The view that looks plausible in the light of recent biology is one that recognizes a gradual evolution of the capacity to experience events. A break between "higher" and "lower" animals is looking less and less likely.

If these continuities are embraced, how might we deny materialism? Panpsychism is one route. This is an extreme option, admitted as such by proponents. People see themselves as pushed to it by standard arguments against materialism and dualism. Those arguments against materialism, again, have adequate replies though I am not addressing them here (see note 18). With those arguments defused, panpsychism loses much of its appeal.

Instead of going further down that road, I'll discuss another alternative to materialism, a kind of "neutral monism." That name is used for a number of views. What I have in mind is not the "Russellian" form, in which physics describes pure structure and hence leaves room for intrinsic properties that are "phenomenal," or close to it. The version I have in mind is seen in Dewey, for example (1929). This view holds that what is basic is neither physical nor mental. The patterns and phenomena we associate with those terms ("mental," "physical") are both manifestations of a more basic reality.

Neutral monism of this kind shades into materialism on one side and panpsychism on the other. It shades into materialism of the more critical kinds, such as what Stoljar calls "o-physicalism" and the *via negativa* physicalism I mentioned earlier. It shades into panpsychism via a "panprotopsychism" of the kind Chalmers has defended (2003).[20] That view holds that what is fundamental is not something mental, but something apt to produce the mental, and something other than the merely physical.

I said earlier that the heart of my view, which I see as materialist, is monism and the construction of mind from things that are not mental. Some kinds of neutral monism would seem to pass this test for being materialist. The mental is not derivative on the paradigmatically physical, but it is derivative, in a sense, on the neutral, and hence on something non-mental. (The paradigmatically physical is derivative in the same way.) This blurring of the distinction between materialism and neutral monism might be OK, or we might look for something stronger to mark out materialism, such as the non-derivative nature of the physical.

In any case, I think there are reasons not believe a view like this. If I understand the positive content of a neutral monist option of this kind, then something that is central is a kind of extreme anti-reductionism, the claim that it is *not at all true* that the paradigmatically physical side of things is an explainer of many other kinds of goings-on.

[20] This kind of neutral monism might also count as o-physicalist in Stoljar's sense. Here is Stoljar's category of an "o-physical" property: "a physical property is a property which *either* is the sort of property required by a complete account of the intrinsic nature of paradigmatic physical objects and their constituents *or* else is a property which metaphysically (or logically) supervenes on the sort of property required by a complete account of the intrinsic nature of paradigmatic physical objects and their constituents" (2001, p. 257). How does panprotopsychism relate to my fairly broad category of the materialist? The closer the hypothesized quasi-mental entities or properties at the bottom are to being mental, the further the view is from one in which the mental is constructed from the non-mental. The details need to be filled in, and some of the distinctions here are probably not sharp.

What we call the "physical" is just the manifestation of some dispositions, a way things behave in some circumstances, and/or a way visible from some points of view.[21]

But it seems very likely that the physical *is* an explainer – a basis, a "ground," as people now say – for *some* of the rest, at least. The physical is basic in some respects, not just a set of behaviors in nature on the same level as others. This is evident in the "easier" cases with respect to the relation between biology and the cognitive side of the mind, especially in the way that the physical and chemical give rise to simpler forms of sensing, memory, and agency.[22]

While neutral monism does make some sense as a rival, it has problems as an overall picture. It denies reasonable kinds of reductionism, in its determination to deny a more contentious kind.

Suppose the overall picture had remained one with "humans versus the rest" – a picture with no gradual evolutionary processes, and humans looking very different from all other creatures. Then I think it would be reasonable to wonder about materialism for longer. More generally, the way to assess materialism is to compare it to other options. I do accept that "none of the above" can be a reasonable move in philosophy; then we know we have to think up something new. If materialism was knocked out, we might say

---

[21] A difference between "Russelian" view and the version of neutral monism that I am discussing here is the fact that the Russellian view preserves the basic scientific part-whole hierarchy in most respects. Physics is, in a sense, at the bottom, but physics, together with the edifice built on it, contains a lacuna with respect to the "phenomenal." In the view I am sketching (which I take to be close to Dewey's late view, and perhaps also the view in William James (1904) that influenced Russell), the status of the standard scientific hierarchy with respect to part-whole relations and the general direction of explanation is revised in a more wholesale way. The physical processes and phenomena that are usually seen as "at the bottom," are not really at this "level," but are just visible in certain contexts – contexts featuring simplicity of interaction and/or contexts involving a particular point of view.

To me, James's own 1904 view seems closer to outright idealism, but if the view is "neutral," then it has this shape. I don't claim that Dewey consistently endorsed this view, but it is pretty clear in *Experience and Nature.*

[22] Earlier I raised the question of the relationship between my broad form of materialism (monism and the non-mental basis of the mental) and this form of neutral monism. The difference between them can be sharpened if it is claimed (as I would claim) that that a complete explanation of the mental has a structure that is in accord with a standard scientific hierarchy with respect to part-whole relations and treats the mental, including the experiential, as a wholly macro-level affair. So we might recognize a basic view (monism and the non-mental explanation of mind) and a strengthening that takes us to materialism as opposed to neutral monism.

that. But materialism is not knocked out. A person can also say: "I can't give any detail yet, but something *in this vicinity* is right." To the extent that the "vicinity" is filled out, there is an alternative view to consider. Otherwise there is not an alternative to consider – nonmaterialist mysterianism is not a worked-out rival. And when we look at the views in other vicinities that have been filled out, we find they have worse problems than materialism.

References

Armstrong, D.M. (1968). *A Materialist Theory of the Mind.* London: Routledge and Kegan Paul.

Armstrong, D.M. (1973). *Belief, Truth, and Knowledge*. Cambridge: Cambridge University Press.

Armstrong, D.M. (1989). *A Combinatorial Theory of Possibility.* Cambridge: Cambridge University Press.

Braddon-Mitchell, D. and Jackson, F. (2006). *Philosophy of Mind and Cognition: An Introduction,* 2nd ed. Hoboken NJ: Wiley-Blackwell.

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

Chalmers, D.J. (2003). "Consciousness and its Place in Nature," in Stich, S.P. & Warfield, T.A. (eds.) *Blackwell Guide to the Philosophy of Mind*, pp. 247–272, Oxford: Blackwell.

Chisholm, R.M. (1951). "Reichenbach on Observing and Perceiving." *Philosophical Studies* 2: 45-48.

Dewey, J. (1929). *Experience and Nature*. Revised edition. Reprinted in J-A Boydston (ed.), *John Dewey: The Later Works, 1925–1953,* Vol. 1. Carbondale: Southern Illinois University Press.

Elliot, E. (2015). *Representation Theorems and the Grounds of Intentionality.* PhD thesis, Australian National University.

Geach, P.T. (1957). *Mental Acts: Their Content and Their Objects.* London: Routledge & Kegan Paul.

Godfrey-Smith, P. (2016) *Other Minds: The Octopus, the Sea, and the Deep Origins of Consciousness*. New York: Farrar, Straus and Giroux.

Godfrey-Smith, P. (2019). "Evolving Across the Explanatory Gap." *Philosophy, Theory, and Practice in Biology*. 11:1.

Hempel, C.G. (1958). "The Theoretician's Dilemma." In *Concepts, Theories, and the Mind-Body Problem,* edited by H. Feigl, M. Scriven, and G. Maxwell. Minnesota Studies in the Philosophy of Science, vol. 2. Minneapolis: University of Minnesota Press.

Jackson, F. (1982). "Epiphenomenal Qualia," *The Philosophical Quarterly* 32: 127-136

Lewis, D.K. (1972). "Psychophysical and Theoretical Identifications." *Australasian Journal of Philosophy* 50: 249-258.

Lewis, D.K. (1994). "Reduction of Mind," In Samuel Guttenplan (ed.), *A Companion to Philosophy of Mind*. Oxford: Blackwell Publishers, pp. 412–431.

Lycan, W.G. (2009). "Giving Dualism its Due," *Australasian Journal of Philosophy* 87: 551–563.

Merker, B. (2005). "The Liabilities of Mobility: A Selection Pressure for the Transition to Consciousness in Animal Evolution," *Consciousness and Cognition* 14: 89–114.

Montero, B. and D. Papineau (2005). "A Defense of the Via Negativa Argument for Physicalism." *Analysis* 65: 233-237

Place, U.T. (1956). "Is Consciousness a Brain Process?" *British Journal of Psychology* 47: 44–50.

Popper, K.R. (1959). *The Logic of Scientific Discovery.* New York: Basic Books.

Psillos, S. (2004). "Ramsey's Ramsey-Sentences." *Vienna Circle Institute Yearbook* 12: 67-90.

Ramsey, F.P. (1926) "Truth and Probability." In Ramsey, *The Foundations of Mathematics and other Logical Essays*, pp. 156-198, edited by R.B. Braithwaite, London: Kegan, Paul, Trench, Trubner & Co. (1999 electronic edition.)

Ramsey, F.R. (1931). "General Propositions and Causality." In *Philosophical Papers*, ed. D. H. Mellor. Cambridge: Cambridge University Press, 1990, pp. 145–63.

Ryle, G. (1949). *The Concept of Mind.* London: Hutchinson.

Smart, J.J.C. (1963). *Philosophy and Scientific Realism.* London: Routledge.

Sober, E. (1990). "Let's Razor Ockham's Razor." *Royal Institute of Philosophy Supplement* 27:73-93

Stoljar, D. (2001). "Two Conceptions of the Physical." *Philosophical and Phenomenological Research* 62: 253-281.