

The Essence of the Mental

Forthcoming in the *European Journal of Philosophy*
Ray Buchanan and Alex Grzankowski

Abstract:

Your belief that Obama is a Democrat wouldn't be the belief that it is if it didn't represent Obama, nor would the pain in your ankle be the state that is if, say, it felt like an itch. Accordingly, it is tempting to hold that phenomenal and representational properties are essential to the mental states that have them. But, as several theorists have forcefully argued (including Kripke (1980) and Burge (1979, 1982)) this attractive idea is seemingly in tension with another equally attractive thesis, namely, the token-identity thesis; the thesis according to which every mental state token is identical with some or other token physical state. In this paper, we show that these seemingly incontrovertible essentialist intuitions are in fact compatible with 'token physicalism' regarding the mental.¹ Given a suitably plentiful ontology of objects, we argue that there *are* physical things with which our token mental states can be identified. This is preferable to existing views that give up the essentiality claims or simply reject the token-identity thesis.

1. Introduction

The phenomenal and representational features of our mental states are not merely contingent marks thereof. Consider your belief that Obama was a Democrat. Could it be the particular state that it is if it represented something else or if it failed to represent at all? Analogously, consider the pain in your foot. Could it exist, but not hurt? Moreover, instantiations of these features, the phenomenal and the representational, are not *merely* necessarily correlated with the existence of our mental states, in the way that, say, instantiating the property of being distinct from Barack Obama is necessarily correlated with

¹ In what follows, we will use 'token identity theory' and 'token physicalism' interchangeably. When we speak of, for example, your belief that Obama is a Democrat, or Rosa-Linda's pain in her foot, we mean to be speaking of token mental states. We will explicitly indicate when we are instead talking about state-types – such as the belief that Obama is a Democrat – when the context does not make it clear already.

your existence. Rather, these features are part of that in virtue of which our mental states are the very states that they are.

One who accepts the foregoing essentialist intuitions about the mental faces an immediate question: *what could a particular mental state of yours be such that it represents, or has a phenomenal character, essentially?* The question is especially pressing for anyone with physicalist sympathies since it is unclear how any particular physical state – for example a token brain state – could have such features *essentially*.

Reflecting on that question, some philosophers have given up the token identity theory (and hence the type identity thesis that entails it) on the grounds that they can't see how to reconcile their view with essentiality intuitions.² In what follows we revisit two arguments – inspired by Kripke (1980) and Burge (1979, 2009) – that purport to show that token physicalism cannot be reconciled with the claims that mental states are essentially representational or phenomenal (Section 1). We then turn to a physicalist-friendly reply that seeks to explain away our essentialist intuitions, holding that phenomenal and representational features are only contingently possessed by the states that have them. We argue that this contingentist retreat fails to successfully explain away the counter-intuitiveness of the view. But, more importantly, the real worry for this retreat is that it is unmotivated since there *is* a way of reconciling token physicalism and the essentialist intuitions (Section 2). Another, more promising, physicalist-friendly reply aims to retain the essentialist intuitions but only at the cost of denying that there are identities holding between the mental and the physical. For example, a physicalist might follow Boyd's (1980) suggestion that our mental states are *constituted* by, but not identical with, some or other physical

² The Burge/Kripke-inspired essentialist arguments against token physicalism don't just threaten the type-identity theory; they also threaten any non-reductive physicalist account on which *actual* mental state tokens are claimed to be identical to physical tokens. For example, consider a non-reductive physicalist who identifies mental state types with second-order functional state-types, but then claims that all of the realizers of those state-types are, in the actual world, themselves physical state tokens, say token brain states. The family of arguments with which we are concerned would threaten this view as well since they purport to show that even our actual mental state tokens cannot be identical to any physical token. Thank you to an anonymous referee for pressing this point.

state (Section 3). Though we are sympathetic with Boyd’s appeal to constitution, and the analogy that it suggests with problems of material constitution, we believe there is a more satisfying physicalist position that retains token identity and all of the reductivist advantages that come along with it (Section 4). By our lights, token physicalists need not worry about the essentialist arguments inspired by Burge, Kripke, and their followers. More specifically, we argue that if one allows that physical states have mental properties – contingently, or otherwise – then physicalists have an easy response to the essentialist worries. As we will see, this is because there is a compelling argument for a plenitudinous ontology that becomes available once one allows in mental properties even contingently and this argument will provide us with additional token entities we can utilize in reply to the Kripke and Burge-inspired arguments (Section 5). To be clear, this is not to say that all physicalist worries are allayed (such as Jackson’s Mary or Chalmers’s Zombies) but we hope we will persuade readers that one family of well known worries can be avoided, given a suitably plenitudinous ontology.

Talk of ‘mental states’ is, of course, ambiguous between state-types and tokens, but in what follows we will focus on tokens.³ There are two reasons for this. The first is that the arguments we consider below are attacks on token-token identity theory.⁴ The second reason is that, *prima facie*, the idea that a mental-state *type* might itself be conscious or represent should seem puzzling. How, for example, could a *type of state* hurt? How could the belief that grass is green, qua abstract type, be about something? But these claims are less mysterious when we focus on the token mental states that fall under them. It’s not that there are no mysteries about how some token state is conscious or represents, but at least it seems like the right kind of thing to be conscious or to represent. In our view, it is mental state tokens and their representational and phenomenal properties that are more fundamental in the order of explanation. We can truthfully speak of

³ Following Goldman (1971) and Kim (1976), we take *events* to be property instantiations and *states* to be perduring events. Our talk of “neural events/states” in what follows should then be understood accordingly.

⁴ Our discussion below could just as easily be couched in terms of token mental *events*, rather than states, without loss.

types of mental states as being conscious or representing only in a derivative sense – insofar as all of their tokens have these properties. Mental-state types can be said to be conscious and represent in much the same way that a type, or kind, of truck might be said to be hard to park. Qua abstract type, the Ford F-150 is not the sort of thing that one can park the way one can park a token truck that takes up space. Hence, the cash-value of the claim that the Ford F-150 (a type of truck) is hard to park is simply that all the tokens of that type are hard to park. So too with mental state-types and their tokens. The cash-value of the claim that the belief that Oscar snores is about Oscar is that all tokens of that type are about Oscar. When it comes to intentionality and the phenomenal, the real action is, in our view, with token states rather than with types.

2. Kripke, Burge, and Token Physicalism

If a mental state token and physical state token differ in their properties, they cannot be identical. There are compelling reasons for thinking that the mental and the physical do so differ. First, consider Kripke's (1980) well known discussion of pain:

Let 'A' name a particular pain sensation, and let 'B' name the corresponding brain state, or the brain state some identity theorist wishes to identify with A. *Prima facie*, it would seem that it is at least logically possible that B should have existed (Jones's brain could have been in exactly that state at the time in question) without Jones feeling any pain at all, and thus without the presence of A. [...] If A and B were identical, the identity would have to be necessary. The difficulty can hardly be evaded by arguing that although B could not exist without A, *being a pain* is merely a contingent property of A, and that therefore the presence of B without pain does not imply the presence of B without A. Can any case of essence be more obvious than the fact that *being a pain* is a necessary property of each pain? (146)

According to Kripke, it is implausible that a particular pain — say, the sensation in your lower back — could have existed without hurting; the painfulness of pains is not plausibly a contingent feature thereof. Burge (2009) offers a structurally similar considerations concerning the representational:

Any given token physical neural event that is a candidate for being identical with a mental event could have been associated with different causal antecedents in the distal environment from those it in fact had. [...] Mental events are different events if they have different representational contents. [...] So the given neural event could exist in a situation in which the mental event with which it is supposed to be identified did not occur, and in which some other mental event (with a different representational content) occurs instead. So the neural event is not identical with the original mental event (234).⁵

As Burge points out, the foregoing argument generalizes, so we ought to conclude that neural events/states are not identical with representational mental events/states. It is debatable how exactly to best formulate these arguments so as to be maximally faithful to the texts, but we think that the following captures the force and spirit of both:⁶

1. All token phenomenal/representational states are essentially phenomenal/representational.
2. No token brain states are essentially phenomenal/representational.
3. So, token phenomenal/representational states are not identical to token brain states.

⁵ See also Burge (1979, 1982). Though Burge focuses on the role of external determinants of content, an analogous argument can be given against most accounts of narrow content, as well. Insofar as a narrow content theorist individuates contentful token states in terms of *non-intrinsic* features, her account is subject to the same worries. It is contingent that a particular brain state has the particular internal functional role that it does.

⁶ Famously, Kripke offers multiple broadly anti-physicalist arguments in *Naming and Necessity*. In what follows we will primarily focus on what we have called The Essentialist Argument that, as we say, we believe captures the spirit of an important line of thought. Below (fn. 26) we will have a bit more to say about another strand of Kripke's arguments against token physicalism.

Call this argument, inspired by Burge and Kripke, ‘The Essentialist Argument’.⁷ There are at least two ways one might then elaborate The Essentialist Argument. First, one might argue that there are no plausible candidates other than brain states for being that to which our token mental states are identical. Second, one might argue that this style of argument will work for any other forthcoming candidates. For instance, everyone might agree that a bit of brain matter in a petri dish feels nothing and represents nothing but might still hold that a structure of firing, embodied neurons may well represent or experience. Still, such an embodied structure *might* not have had any phenomenal or representational features. One might think that this style of attack will apply to any physicalist candidate we come up with. If The Essentialist Argument is persuasive (as we think it is), physicalism about the mental looks to be in trouble.⁸

3. The Contingentist Retreat

There have been a number of physicalist reactions to this style of argument. Here we consider an attempt to deny Premise 1, originally due to Feldman (1974).⁹

According to Feldman, although there is a nearby *de dicto* modal truth, 1 is simply false when read *de re*. For concreteness, consider the following which looks to follow from the more general 1 as stated above: It is essentially the case that all pains are painful ([essentially] $\forall x (\text{Pain}(x) \supset \text{Painful}(x))$). This seemingly analytic, *de dicto* truth (the reply continues) provides an explanation of why we are apt to judge premise 1 true. However, the corresponding *de re* modal claim regarding tokens is false: $\forall x (\text{Pain}(x) \supset [\text{essentially}]$

⁷ For further discussion of Kripke’s argument see Feldman (1974) (more anon); for Burge, see Kobes (2009).

⁸ It is worth highlighting that The Essentialist Argument is distinct from other, familiar attacks on reductive physicalism such as the conceivability of zombies and the Knowledge Argument. Even if everything we argue for in what follows concerning the Essentiality argument is correct, the physicalists must still face these distinct arguments.

⁹ Feldman’s presentation of Kripke’s argument differs slightly from the more stream-lined version we presented in Section 1. For present purposes, however, these differences will be un-important.

Painful(x)). That is, when it comes to any particular token pain event, *it* might not have hurt. As Feldman elaborates:

[A]ny serious materialist should recognize that his view entails that painfulness is never part of the essence of a pain-event. Pain-events are experienced as they are only as a result of contingent laws of nature. Early materialists may have had just this point in mind when they asserted, somewhat confusedly in my opinion, that certain brain-events are “contingently identical to” pain-events. Understood charitably, what they must have meant is that it is contingently true that brain-events of certain descriptions are identical to pain-events. These very brain-events, had the laws of nature been different, would of course still have been self-identical, but would not have been identical to anything that would, under those circumstances, have been a pain-event. Thus, such events are not essentially painful (675).

For Feldman, a token pain is a brain event, yet it is contingent that it is felt as a pain. The above concerns pain, but one can see how to tell the same sort of story for representational states drawing on Burge.

Recall that for Burge, a brain state represents in virtue of its relations to extra-cranial environmental facts.

Hence, it is a contingent matter whether a brain state is related in such ways to the environment of the thinker so as to represent at all. A physicalist might then be tempted to follow Feldman and claim that while it is true that [essentially] $\forall x$ (The Belief That Oscar Snores(x) \supset About Oscar(x)), it is false that it is essential to, say, your token belief that Oscar snores that it is essentially about Oscar, or anything else.¹⁰

So, following Feldman on pain, a physicalist/materialist should deny that representing is of the essence of a token representational state.

¹⁰ See Gibbons (1993) and Frances (2007) for additional discussion and defense of this contingentist retreat.

From a certain vantage point, this is initially a compelling physicalist reply to Kripke and Burge. We agree that any given token brain state that happens, here and now, to hurt, might have existed but not hurt. Likewise, any token brain state that happens to represent grass as being green, might have represented something else or nothing at all. Nevertheless, there are concerns. Contra Feldman, we are doubtful that the debate crucially depends on any would-be scope ambiguity, since the Burgean/Kripkean arguments can be trivially recast so as to render any appeal to scope useless: we can simply refocus the debate on the *de re* claims regarding our token *mental* states. We might ask whether, say, Oscar's belief that Carla snores is such that *it* is essentially about Carla, or whether Oscar's pain is such that *it* might not have hurt. When we focus our attention in this way, we hope that these claims seem as manifestly false to you as they do to us. We grant that *if* the foregoing essence claim about Carla's belief were really about a brain state, it would be false; but we suspect that for most competent speakers, it would take nothing short of 'bribery, threats, [or] hypnosis' to convince them this is what they literally meant.¹¹ This is, we think, some initial evidence for thinking that Feldman's gloss on the *de re* essence claims is incorrect and brain states aren't the best candidates for what we are talking about when we speak of token phenomenal or representational states.

But recall where we began: what could a particular mental state of yours *be* such that it represents, or has a phenomenal character, essentially? There is, here, a putative data point that it would be nice to capture if possible. The premises seem intuitively compelling, so much so that we think that a reply that accommodates the essentialist intuitions is to be preferred. Now, if one thinks it is *just obvious* that, say, Jen's belief is a brain state of hers, then we see why a physicalist ought to follow Feldman and give up the view that there is anything that is both Jen's belief and that represents essentially. But we think (and we think Kripke and Burge must think) that this simply *isn't* obvious. Moreover, as we will argue below, there *are* other candidate things with which we can identify token mental states and this will allow us to

¹¹ The phrase is borrowed from Richard's (1990, p.125) discussion of Millian theories of belief reports.

accommodate the compelling premises. The contingentist reply is plausible if you have *already made up your mind that mental-state tokens must be brain-state tokens*, but for those who think it remains an open question, this style of reply gets us ahead of ourselves. We should ask, even as physicalists, are there any other candidate things which are both physicalistically acceptable and that essentially represent or have phenomenal properties? If the only sensible view available is that token mental states are identical to token brain states, then one should adopt Feldman's approach and learn to live with the rejection of the essentialist intuitions. But wouldn't it be preferable if we could somehow hold onto the essentialist intuitions without giving up token physicalism?

4. A Different Tack: Constitution

There is a second reply to The Essentialist Argument, due to Boyd (1980), to which we are much more sympathetic.¹²

Boyd argues that physicalists can accept the conclusion of The Essentialist Argument, but deny that it undermines their view. According to Boyd, physicalists can grant that although our mental states are not identical to physical states, they are *constituted* by them. Physicalists do not need token identities, only constitution. With respect to token physicalism, Boyd (1980) writes:

[I]t is a mistake to understand materialism as entailing that each token mental event, state, or process is identical to some quite specific molecular or physiological event, state, or process. (99) [...] The issue is not identity versus correlation [as the dualist suggests], but composition versus correlation.¹³ The issue is whether the physical state associated with a

¹² See also Rudder Baker (2000) and Tye (2003) for discussion of the view that mental states and persons are constituted by, but not identical with, the physical.

¹³ In *Naming and Necessity*, Kripke briefly considers, and sets aside for future discussion, an analogous view to the effect that persons are constituted by, but not identical with, their bodies (fn 74, p. 145).

mental state constitutes or realizes the mental state in question, or whether, on the contrary, it merely correlates with it. (102)

On Boyd's view, it is false that Jen's belief that Oscar snores is identical to any brain state, but, if physicalism is true, it is nevertheless *constituted* by something physical.

We are sympathetic to Boyd's claim that physical tokens constitute mental tokens and we agree The Essentialist Argument is moving and pressing. But we are ultimately unsatisfied with Boyd's reply. Our question remains: what could a particular mental state of yours *be* such that it represents, or has a phenomenal character, essentially? Boyd has not answered this residual question. And notice that the question does not simply smuggle in the 'is' of constitution; we are perfectly happy to paraphrase: what could a particular mental state of yours *be identical to* such that it represents, or has a phenomenal character, essentially? A token physicalist ought to be able to tell us which physical thing a given mental state is identical to and if there is no answer, we think dualists should feel, if not vindicated, certainly motivated. This thought can be brought into sharper relief with a comparison. Consider someone who claims that a lump of clay constitutes a statue but is not identical to it and who yet affirms physicalism. Does not a pressing question remain, namely, to which physical thing is the statue identical? A: 'The statue is composed of the physical matter, but it is not identical to it.' B: 'What, then, in your physicalist worldview, *is* the statue?' B's question to A is a perfectly good question that should, in our view, not be evaded. And stopping at the answer 'it is identical to itself and no other thing' is unsatisfying if we seek to support physicalism. The statue is certainly something and moreover, if physicalism is true, it is something physical. Which thing?¹⁴

¹⁴ Perhaps a proponent of Boyd's constitution-based approach would resist our question above regarding the identity of the constituted mental particular. One could dig in their heels just here, but in reply we think even the fans of constitution should feel some pressure to answer our question. An analogy helps here. Suppose one holds that Goliath is constituted, but not identical to, Lumpl. But, if so, how could the question of 'so which thing *is* Goliath, then?' not be answerable? The idea that this isn't an admissible question, or is a question that can receive no answer other than 'it is itself', strikes us as unsatisfactory. Constitution connects the statue to the clay, but questions of identity don't thereby dissolve. In fact, as we

Boyd himself seems tempted to agree. His aim is to show that physicalists can stop once they have shown to their satisfaction that mental states are constituted by the physical, but he flirts with an identification:

These considerations have the effect of making token events, states, and processes seem less like stereotypical “individuals” and more like type events, states, or processes – more like “universals” – in that a token event, for example, may have more than one instance (although in different possible worlds), to none of which it need be identical. (102)

In effect, the view of mental tokens that Boyd is driven to holds that even they are multiply realizable. We will come back to that idea in the next paragraph, but for now, it is notable that Boyd bothers to try out an answer to our identity question. This lends credence to our claim that the question cries out for an answer.

Briefly, let us mention why we aren't happy to adopt Boyd's suggestion about mental tokens. The idea that a token mental state might itself be multiply realizable, and, hence, more like a universal, raises more questions than it answers. For example, if Jen's belief that Oscar snores is a type-level entity, how could it non-derivatively represent Oscar, much less essentially so? Likewise, if the pain in Bill's lower back is a universal, how can *it* hurt? As we noted earlier, we find it mysterious how universals or types might have the relevant properties we are concerned with if not by virtue of their tokens having those properties. But if tokens themselves are thought of as types or type-like things, we can't see how the tokens might have the sorts of properties they are alleged to have.

The view we wish to offer takes on board much of what Boyd advocates, but we believe we are in a position to answer our question: which thing is identical to the mental state in question?

5 Plenitudinous Token Physicalism

will show in the main text, Boyd himself seems to feel the pull of the question about identity. Thanks to an anonymous referee for discussion.

In the literature spawned by Burge and Kripke's essentialist arguments, it has occasionally been suggested that physicalists need a broader, more liberal conception of what mental state tokens might be; token neural structures might not be the physicalists's best bet.¹⁵ We think this suspicion is correct. As we are about to argue, the physicalist has available better (and independently motivated) options.

The Essentialist Argument presupposes that some things have properties essentially and others contingently. But it has been powerfully argued that proper appreciation of this distinction leads to a 'plenitudinous' ontology. To appreciate why, let's momentarily shift gears and consider an argument for universalism about composition. Consider the following example from Van Cleve (2008):

"[I]magine a hook, a line, and a pole that I buy in order to make a fishing rod. When the three components lie in separate bins at the hardware store, they do not (according to likely moderate principles) compose any further object. Side by side on my workbench, they still do not compose a new object. When the line is tied tightly enough around the hook at one end and the pole at the other, they *do* compose a new object. But what happens before the knots are tight? What if the diameter of the knots is several times that of the pole, so that lifting the pole leaves the line behind? What if the diameter is now one millimeter less? You get the idea: to avoid objectionable arbitrariness or indeterminacy, we must say that the three components *always* [or never] composed a further thing." (17)

Common sense ontology allows hooks, lines, poles and complete fishing rods and one might hope to avoid more entities than those. But what the example shows us is that it is exceedingly difficult to see how any cutoff could be non-arbitrary. The reasons one might offer in support of including the completed rod in their ontology apply to the nearest neighboring case. And now we are off to the races, seeking some substantive difference between neighbors that are as similar as piles of sand differing by one grain.

¹⁵ See, for instance, McGinn (1989) and Nagel (1965). McGinn responds to Burge by identifying token representational states not with neural states/events but rather with neural states in tandem with their relational, historical features. A token belief that Fido snores is, hence, quite literally spread out over space and time. See Macdonald (1990) for reasons for resisting McGinn. On the view offered below, these unpalatable consequences are avoided.

According to the Universalist, it would be arbitrary, non-principled, and *ad hoc* to allow the completed rod but deny the other candidates. So we ought, the Universalist claims, rules them all in.¹⁶

We wish to marshal a similar argument, not for Universalism but rather for a plentitude of coincident entities. Consider an example due to Leslie (2011). Suppose there is an object with exactly five properties, all of which are modally independent from one another. Suppose further that we have satisfied ourselves that the entity in question has the first two properties essentially and the other three contingently. On what grounds might we deny that there is another, coincident entity that has three of the properties essentially and two contingently? And similarly for other combinations. Once one has adopted a view according to which some properties are had essentially and others contingently, it is difficult to see what principled reason could be given for allowing some combinations into one's ontology while ruling out others. Now, since some properties are not modally independent from one another (such as being blue and being colored), not just anything whatsoever goes.¹⁷ But, whatever the relevant combinatorial restrictions might be, we will be left with a vast ontology of coincident entities.¹⁸

Returning to The Essentialist Argument, recall that our strategy is to admit that the argument is a good one but to show that physicalists needn't be deeply concerned. For the reasons just given concerning the

¹⁶ As van Cleve (*ibid.*) notes, another reaction to the argument is to endorse nihilism. Given the dialectic in which we are engaged, this is a non-starter. We are concerned with theorists such as Burge and Kripke who hold that there are brains, neurons, and so on. In light of this, the argument from arbitrary cutoffs takes one to the maximal view rather than the minimal view.

¹⁷ Exactly how to restrict the plentitude of entities is a matter of contemporary debate. See Fairchild (2019, 2020), Kriegel (2019), Leslie (2011), and Yablo (1987) for further discussion. We follow Leslie by focusing on 'strongly modally independent' properties. Let a *modal profile* of an entity be a partition of properties into those had essentially and accidentally. An admissible modal profile M for an entity x will then comprise those properties such that it is possible for x to have any subset of M while lacking the others. As noted in the main text, nothing can be both essentially blue and accidentally colored since such a partition is inconsistent on metaphysical grounds (being colored and being blue aren't, to use Leslie's phrase, 'strongly modally independent'). We see no reason for thinking these complications will affect anything of substance regarding our discussion in the body of the paper.

¹⁸ See also Fine (1999); Hawthorne (2006); Johnston (2006); Yablo (1987). Note that while Universalism holds that there are many partially overlapping entities, "many thingism" holds that there are many co-located entities differing in having properties essentially or contingently. For a general discussion and introduction see Fairchild (2020). See Kriegel (2019) for a discussion of some of the complications many thingism leads to in the theory of reference.

object with exactly five properties, it's hard to see how to accept the first premise of The Essentialist Argument without accepting a plenitudinous ontology of coincident entities. But, provided there are physical token entities that even contingently have mental features, then there *are* (in light of the foregoing plenitudinous arguments) physical token entities that essentially represent and essentially hurt. For example, consider your belief that grass is green and a brain state, call it 'b'. Call the property that a would-be physicalist claims to suffice for representation/phenomenal character 'F' (perhaps F is the property of having a certain functional role or such and such tracking feature).¹⁹ As Burge and Kripke would seemingly agree, b will have F contingently, if at all. But, notice that for those sympathetic to the possibility of essential properties, brain state b will also have various other properties essentially (*e.g.* being made of biological stuff). For illustration, pretend there is just one such property and call it 'G'. According to the plenitudinarian there is some *other* entity, distinct from b that contingently has G but essentially has F. That entity, we submit, is a good candidate for being identical to your token mental state. And since the 'stuff' in question is physical and the properties are physicalistically acceptable, this is an entirely physicalistically acceptable entity.²⁰

Those who find such 'abundant' ontologies implausible may feel unsatisfied, so let's pause to remind ourselves how we got here: Anyone who finds The Essentialist Argument compelling must take essential

¹⁹ Some dualists argue that there are no physicalistically acceptable properties that suffice for mentality. For example, the Zombie Argument purports to show that functional duplication does not guarantee phenomenal duplication. It is important to flag that we are not presently engaged with that concern. Our aim is to diffuse a different attack on physicalism, namely one that purports to show that token identity theory fails.

²⁰ In the vast literature on functionalism it has been, at least once, suggested that mental state tokens should not be identified either with the neural realiser of a role nor the second-order role itself. Consider the view that McLaughlin (2006, 2007) dubs 'Physical Realization-Role-Functionalism'; a view according to which "no mental event token is a physical event token, but every mental event token is a functional event token that is realized by a physical event" (45). See also Shoemaker (2001) for an interesting precedent.. In effect, this is a particular instance of the general kind of view that we offer in the main text. For the PR-functionalist, a mental state token might be identified with a physical neural token *qua* realiser of the relevant functional role. Though we are sympathetic to this specific suggestion, it is important to note that it was not offered as a response to The Essentialist Argument. Moreover, the response we offer to The Essentialist Argument is much more general. One need not be a functionalist of any stripe about the representational or the phenomenal to endorse our response.

properties seriously. If you don't, you should deny premise 1 immediately. But if you are willing to grant this reasonable premise, then for entirely independent reasons, you ought to be plentiful in your ontology. Once this is appreciated, token physicalists have a lot more than token brain states to work with! Anyone warm to the idea that there are physical things that even accidentally instantiate mental properties should be willing to admit, on the basis of the argument for a plentitude, that there are further entities that essentially have the mental properties.

At this juncture one might reasonably ask, but what *is* one's (say) pain on this picture? At a very general level, the answer is that it is a physical entity individuated (at least in part) by its essential properties. Plausibly, the physical entity is to be found in the head and has the property of hurting essentially. There are various ways that one might articulate this theory of objects, but one way that we find attractive is advanced by Fine (1982). According to Fine, our ontology should include 'qua-objects', objects that are, by their very nature, fusions of matter and their essential properties.²¹ Qua-objects obey a few simple principles:

- a qua F exists iff a and F exist and a has F at some time.
- a qua F exists at time t iff a qua F exists, a exists at t and a has F at t.
- a qua F exists is located at position p at time t iff a qua F exists at t and a is located at position p at t.
- The object x is a part of a qua F iff $x=a$ or $x=F$ or x is a part of a or x is a part of F.²²

²¹ The view is recognisably Neo-Aristotelian and one could just as easily put the view in terms of "matter and form" which Fine himself utilizes in more recent work (2003, 2008).

²² We are making an important simplifying assumption in our characterization of qua objects, namely that the matter of *a*, the "basis", does not change. Fine (1982) sketches a plausible way we might drop the non-changing-matter simplification: "In case the matter is allowed to change, we should first take any matter that ever constitutes the statue; gloss it with the property of having the Goliath-shape; restrict the resulting qua object to those world-times at which the matter constitutes the statue; and then aggregate all of the resulting segments of qua objects. The form of other material things will be similar, though perhaps with greater complexity in the structure of the basis and the gloss." (101). Also see fn. 25 below.

On such a view, many of the physical objects we are surrounded by are in fact qua-objects. For example, there exists an array of densely organized particles at University College London upon which an auto-
icon of Bentham is perched. For familiar reasons, many have held that those particles are non-identical with a chair – the particles can survive smashing but the chair can't, for example. Those particular particles might have coalesced in that very location, time, and shape entirely by accident. If they had, they would be very well suited for sitting but would not thereby be a chair. Chairs are (at least in part) things designed for sitting. But which thing, precisely, *is* the chair? The chair (at time *t*) is the precise array (at *t*) qua chair-shaped-thing-designed-for-sitting (or whatever property the best theory of chairs delivers). The chair massively overlaps with the array but, unlike the array, it is essentially designed for sitting.²³

Another example might be helpful. Consider the inscription 'cat'. It represents cats but it could have represented tables, dogs, or nothing at all. But consider that inscription in a certain light: Fix any baptismal events and fix all the usage facts present and past; now ask, does 'cat' qua inscription-with-its-history-and-usage essentially represent cats? We think it does. If all of the sorts of things that make it the case that a word represents what it does are held fixed and if the object in question is then conceived of as something, by its very nature, tethered to those properties, it is hard to see how it could fail to represent (barring the further addition of properties that somehow disrupt the properties in question). As usually conceived, the inscription 'cat' essentially has, say, a shape but merely contingently represents cats. But there should be no bar on considering the distinct object that essentially represents cats. Given that we have some sense of the properties that contingently make it the case that 'cat' represents cats, we have

²³ One might wonder why we should accept that qua objects *really* exist, rather than claiming that they merely correspond to different possible ways of describing their respective material bases. In the present dialectic, our reason is this: there are entities that, as a *de re* matter, seem to have certain properties that are not shared by their underlying physical basis. For example, a mental state token *m* can (seemingly) have a property – say, *essentially hurting* – that the relevant underlying neural state *n* does not. By the Indiscernibility of Identicals, however, this suggests that $m \neq n$. See Fine (2003) for a critical discussion of this descriptivist alternative.

some purchase on the properties that are had by the distinct object – ‘cat’ qua inscription-with-its-history-and-usage – that are essential to it.²⁴

Returning to the mental, our proposal is that token mental states are qua-objects. Suppose that Jen believes that Oscar snores. Jen’s belief is essentially about Oscar. Jen is also in a neural state, *n*, that we might have initially thought was a good candidate for being identical to her belief. It is not, however, since her belief represents essentially but *n* doesn’t. But consider the object importantly related to *n*, call it *n*-qua-*F* (where *F* is a property such as having the function of tracking Oscar’s snoring, or whatever physicalist-friendly property might account for intentionality).²⁵ This firing-neural-cluster-qua-*F* will abide by the principles laid out above. Of particular importance, *n*-qua-*F*, by its very nature, has *F*. But on the physicalist theory of intentionality, to have *F* just is to represent that Oscar snores. So *n*-qua-*F*, which is distinct from *n*, looks to be exactly the right entity with which to identify Jen’s belief. And there is a recipe for applying this reasoning to pains or any other mental state tokens: (i) locate the physicalist’s favorite property, call it ‘*F*’, to which the mental property in question is reduced (for example, a representationalist might hold that being in pain is being in a state that represents tissue damage); (ii)

²⁴ For present purposes we hope the foregoing sketch suffices. For a detailed defense of qua-objects see Fine (1982, 2003, 2008).

²⁵ Two important qualifications. First, some physicalists might prefer to say that it is not the neural state token but rather the neurons that compose that state that have the function of tracking Oscar. For now, the crucial point is that as long as there is some physicalistically acceptable thing – be it the neural state, or the neurons in that state – that is a plausible candidate for having features such as tracking such and such state of affairs, or such and such functional role, the main points of our subsequent discussion can be accepted.

Secondly, the sketch just given in the body of the paper must be complicated if we are to allow that a particular mental state token could potentially survive the loss of, say, a particular electron. Following Fine (see fn. 22 above), we are optimistic that it can be plausibly elaborated so as to allow for such a possibility: in particular, if we identify a perduring token mental state with the aggregate of temporal slices of distinct qua-objects we can then make room for accommodating change and persistence through temporal and modal space. This issue is an absolutely crucial one for theorists like us who endorse qua-objects, but we are optimistic that Fine’s (1982) suggestion can be made good on. Showing that this is so must, however, wait for another occasion.

locate the best candidate object, call it ‘a’, that contingently has F (presumably a neural state but follow the best theory); a-qua-F will be identical to the particular mental state in question.²⁶

For those who wish to save token physicalism from The Essentiality Argument, we think the foregoing provides an attractive path. Our favored response is of a conditional form: if physical states have mental features at all (even if merely contingently), then there will be physicalistically acceptable entities that essentially have such features, hence satisfying Kripke and Burge’s essentialist demands. We have not directly argued for the antecedent of this conditional, but we think it is imminently plausible (notice that even most contemporary dualists would accept it). After all, if the proponent of the essentialist worries were to deny it, they would be committing to a variety of substance dualism that brings with it familiar worries.²⁷ Why is this so? To hold that physical states have mental properties neither essentially nor contingently is to deny that they have mental properties at all. It would follow that whatever it is that is a bearer of mental properties cannot itself be a physical thing. We hope that embracing the antecedent of the conditional above and our argument from it will strike one as more plausible than a return to a

²⁶ An interesting consequence of our view is that mental state tokens cannot exist without their underlying neural/physical bases. This consequence is in direct tension with an alternative argument that Kripke considers against the token identity theory. In discussing what he calls the ‘converse problem’ for the token identity theory (ibid., 147), Kripke writes that the identity theory fails since, for example, it might initially seem possible that a particular token pain could have existed without the underlying brain state. As he suggests, however, a physicalist might argue for an error theory regarding this ‘Cartesian intuition’ and indeed we think they can. We think that any plausible version of the identity theory should allow that numerically distinct mental state tokens could have exactly the same phenomenal character. Once this is allowed, it is no surprise that one might be initially tempted to claim that, say, Jen’s pain in her foot, P1, could have existed without the underlying brain state. What we imagine is simply a different pain that has the same phenomenal character. Though P1 could not have existed without its underlying physical base, a distinct, though qualitatively identical, pain could have. As the discussion in the body of the paper should make clear, we do not think the initial intuition that underwrites The Essentialist Argument – *i.e.*, that mental state tokens essentially have the phenomenal/representational features that they do – can be explained away so easily. The error theory just sketched for Kripke’s ‘converse problem’ is obviously of no help in explaining away why it seems possible for the neural/physical token to exist without instantiating any particular mental property.

²⁷ See, for example, Smullyan (2002).

dualism of substances.²⁸ If the foregoing is correct, we will have diffused one important challenge facing many physicalists.²⁹

One might wonder whether our own view is itself a dualism of sorts (or, really, a many-ism). After all, where you might have thought there was but one thing, there are many things. This is correct, but the view should be in no way worrisome for a physicalist. The entity identical to your mental state is not non-located and it is not, even in part, located anywhere other than where your brain is located.³⁰ The object a-qua-F is located exactly where a is. Moreover, there are no reasons for thinking that the view leads to an overabundance of causes. To see this, compare again with the chair and array of particles. Bentham's chair is non-identical to the array that constitutes it. The chair weighs, say, eight pounds. The array does too. One might then worry, shouldn't these two items together weigh sixteen? But this sort of concern is diffused if we consider the *shared* properties of the chair and the array.³¹ Although only the chair has the property of being designed essentially, *both* have the property of being located beneath Bentham, *both* have the property of weighing eight pounds, *both* have the property of being brown, and so on. The chair and the array instantiate many properties and many of those instances are shared instances. The chair's brownness, for example, is the array's brownness. Provided that the causally efficacious property instances of both are shared, there is no threat of an absurd overabundance of causes.³² Similarly, the properties of your belief or your pain that are causally relevant are properties such as firing at such and such Hz, having such and such mass, and so on. These are property instances that your mental state (which is identical to a qua-object) shares with the more familiar object, your brain. So,

²⁸ See Lycan (2009) and Schneider (2012) for some important recent discussion of substance dualism versus property dualism.

²⁹ See footnote 2.

³⁰ See McGinn (1989) and Frances (2007).

³¹ See Barker and Jago (2018), Jago (2018), and Paul (2007) for discussion. For a discussion of closely related issues as applied directly to functionalism and realization, see Wilson (2011).

³² For further discussion and an alternative in terms of counterparts, see Tiehen (2019).

brainstates and mental states needn't be in causal competition. A belief isn't identical to any brain state, but they do massively overlap in their property instances.^{33, 34}

Bibliography

Barker, Stephen and Jago, Mark (2018). 'Material objects and essential bundle theory'.

Philosophical Studies 175 (12):2969-2986.

Baker, Lynne Rudder (2000). *Persons and Bodies: A Constitution View*. Cambridge University Press.

Boyd, Richard (1980). 'Materialism without reductionism: What physicalism does not entail'.

In Ned Block (ed.), *Readings in the Philosophy of Psychology*. , Vol 1. pp. 1--67.

Burge, Tyler (1979). Individualism and the mental. *Midwest Studies in Philosophy* 4 (1):73-122.

Burge, Tyler (1982). Other bodies. In Andrew Woodfield (ed.), *Thought and Object*. Oxford University Press.

Fairchild, Maegan (2019). 'The Barest Flutter of the Smallest Leaf: Understanding Material Plenitude.' *Philosophical Review* 128 (2):143-178.

Fairchild, Maegan (2020). Varieties of plenitude. *Philosophy Compass* 15 (3).

Feldman, Fred (1974). 'Kripke on the identity theory'. *Journal of Philosophy* 71 (October):665-76.

³³ If a token mental state M is a qua-object whose 'base' is itself a physicalistically acceptable entity (e.g., a neural token) it should be no mystery how M itself can play a causal/explanatory role. There is, however, a residual worry: does the neural component of M, call it n, leave any distinctive causal work for the non-neural "gloss" of n? In this way, one might worry that a kind of epiphenomenalism looms. Establishing whether this is so (and moreover whether it is an objectionable form of epiphenomenalism) will have to wait for another occasion.

³⁴ Acknowledgements: For helpful comments, feedback, and discussion, we would like to thank an anonymous referee for this journal, as well as Ali Boyle, Mark Balaguer, Laurenz Cassar, Zoe Jenkin, Jon Litland, Megan Fairchild, Craig French, Mike Martin, Cynthia McDonald, David Papineau, Jake Quity-Dunn, Michael Tye, Wayne Wu, and audiences at San Raffaele's Mark of the Mental Research Project and the University of Manchester.

- Fine, Kit (1982). 'Acts, Events and Things Language and Ontology'. Proceedings of the 6th Wittgenstein Symposium.
- Fine, Kit (2003). 'The non-identity of a material thing and its matter'. *Mind* 112 (446):195-234.
- Fine, Kit (2008). 'Coincidence and Form'. *Aristotelian Society Supplementary Volume* 82 (1):101-118.
- Frances, Bryan (2007). 'Externalism, Physicalism, Statues, and Hunks'. *Philosophical Studies* 133 (2):199-232.
- Gibbons, John (1993). 'Identity Without Supervenience'. *Philosophical Studies* 70 (1):59-79.
- Goldman, Alvin I. (1971). The individuation of action. *Journal of Philosophy* 68 (21):761-774.
- Jago, Mark (2018). 'Essential bundle theory and modality'. *Synthese*:1-16.
- Kim, Jaegwon (1976). Events as Property Exemplifications. In M. Brand & D. Walton (eds.), *Action Theory*. D. Reidel. pp. 310-326.
- Kobes, Bernard W. (2009). 'Burge's dualism'. In Robert C. Koons & George Bealer (eds.), *The Waning of Materialism: New Essays*. Oxford University Press.
- Koons, Robert C. & Bealer, George (eds.) (2009). *The Waning of Materialism*. Oxford University Press.
- Kriegel, U., 2019. Essentialist Plenitude and the Semantics of Proper Names. *Metaphysics*, 2(1), pp.16–25. DOI: <http://doi.org/10.5334/met.19>
- Kripke, Saul A. (1980). *Naming and Necessity*. Harvard University Press.
- Leslie, Sarah-Jane (2011). 'Essence, plenitude, and paradox'. *Philosophical Perspectives* 25 (1):277-296.

- Loar, Brian (1981). *Mind and Meaning*. Cambridge University Press.
- Lycan, William G. (2009). Giving dualism its due. *Australasian Journal of Philosophy* 87 (4):551-563.
- MacDonald, C. (1990). Weak externalism and mind-body identity. *Mind* 99 (395):387-404.
- McGinn, Colin (1989). *Mental Content*. Blackwell.
- McLaughlin, Brian P. (2007). Mental causation and Shoemaker-realization. *Erkenntnis* 67 (2):149 - 172.
- McLaughlin, Brian P. (2006). Is role-functionalism committed to epiphenomenalism? *Journal of Consciousness Studies* 13 (1-2):39-66.
- Nagel, Thomas (1965). Physicalism. *Philosophical Review* 74 (July):339-56.
- Paul, L. A. (2007). 'Constitutive Overdetermination;'. In J. K. Campbell, M. O'Rourke & H. S. Silverstein (eds.), *Causation and Explanation*. MIT Press. pp. 265-290.
- Richard, Mark (1990). *Propositional Attitudes*. New York: Cambridge University Press.
- Shoemaker, S. (2001). Realization and mental causation. In G. Gillett & B. Loewer (Eds.), *Physicalism and its discontents* (pp. 74-98). Cambridge: Cambridge University Press.
- Smullyan, Raymond M. (2002). An Unfortunate Dualist. In David J. Chalmers (ed.), *Philosophy of Mind: Classical and Contemporary Readings*. Oup Usa.
- Schneider, Susan (2012). Why property dualists must reject substance physicalism. *Philosophical Studies* 157 (1):61-76.
- Tiehen, Justin (2019). 'How Counterpart Theory Saves Nonreductive Physicalism'. *Mind* 128 (509):139-174.
- Tye, Michael (2003). *Consciousness and Persons: Unity and Identity*. MIT Press.
- Van Cleve, James (2008). The moon and sixpence : a defense of mereological universalism. In

Theodore Sider, John Hawthorne & Dean W. Zimmerman (eds.), *Contemporary Debates in Metaphysics*. Blackwell.

Wilson, Jessica M. (2011). Non-reductive realization and the powers-based subset strategy.

The Monist (Issue on Powers) 94 (1):121-154.

Yablo, Stephen. (1987). 'Identity, Essence, and Indiscernibility'. *Journal of Philosophy*, 84, 293–314.