Strange Experience: Why Experience Without Access Makes No Sense

Jackson Kernion

1 Introduction: A Strange Pain

Philosophers regularly commit themselves to a conceptual distinction between phenomenal experience and functional structure. As the thought goes, you can't learn anything about an organism's cognitive architecture just by learning about whether/what that organism feels, and vice versa.

But this is a mistake. It generates a series of unsatisfying, intractable debates.

To see why this is a mistake, I'll be considering test cases like the following:

Imagine that you're experiencing incredible pain right now: a sharp, piercing pain in your lower back. But there's something strange about this pain. It doesn't feature in your mental life in the way that 'normal'[1] experiences do. It's causally isolated from your other 'conscious' mental processes: you're unable to determine the pain's location (even though you *feel* it in your lower back), you can't describe its qualities (although it *feels* sharp and piercing), and you're unable to notice that your back feels any different from how it normally feels (even though your back does *feel* very different from the way it normally feels). This pain could continue, and you could go about your day as you normally would without it 'making a difference' or distracting you. You could even sit down to read this paper and think "Wow, what a horrible scenario!" without realizing that this describes precisely your own situation.

This strikes me as a strange sort of pain—so strange that I doubt we can actually make sense of it.

It's not the absence of pain *behavior* (or any outwardly observable effects) that makes the pain strange. Instead, what's strange is that the pain doesn't show up to *you*: it has no 'inwardly' observable effects. You're not merely doing a great job of masking your pain from others; rather, your pain is, in a sense, masked from *you*.

[1] I'll use single quotes throughout this introduction for seemingly natural—but potentially misleading—ways of describing the scenario in question.

Of course, this isolated pain wouldn't seem so strange were we to conceive of it as not really yours, as belonging to some *other* conscious mind. But that's not the case I've described. This is what's so strange: this pain is in one sense *yours* (as something that you're experiencing) and in another sense *not yours* (as something functionally isolated and 'removed' from your other mental states).

Sustained reflection will show that this central tension cannot be resolved. This is instructive: it tells us that having a pain—or *any* kind of experience—necessarily involves having a certain kind of functional 'access' to it. Without meeting such a functional constraint, we can't make sense of experiences *as experiences*.

This upends the conventional wisdom that our concept of conscious experience is fundamentally distinct from anything functional. This is significant, since such conventional wisdom plays a central role in shaping contemporary debates in philosophy of mind. For instance, the debate over whether phenomenology overflows cognitive access[2] makes sense only if there's a real, live question about if and how experience and access can come apart. But they can't. And so the insistence that they *can* come apart has left us looking for a kind of empirical support that is, in principle, impossible to have.

That's but one example. I'll leave the discussion of other downstream consequences for another time. My focus in this paper will be showing how the conventional view is untenable in the face of a particular class of cases involving strange experiences.

The plan for the rest of the paper: In section 2, I lay out in more precise terms the conventional view I'm targeting. In section 3, I expand the strange pain case into a more robust skeptical challenge. In section 4, I work through some strategies for pushing back against this challenge, showing why none of them work. In a final section, I note some key features of the view we should adopt in light of the conventional view's failure.

2 Conceptual Dualism

2.1 Two Ways of Thinking about Consciousness

To rehearse a common refrain, we have two different ways of thinking about 'conscious' mental states.

First, we can think of mental states in phenomenal terms: in terms of *what it feels like*[3] to be in that state, in terms of its *subjective qualitative character*. To

[2] Block (2011) got the debate going. For a nice overview of how different empirical methodologies have failed to settle the issue, see Phillips (2018). Other important entries include XXXX.

[3] Nagel (1974)

think of a mental state in these terms is to think of it as *phenomenally (p-) conscious*[4].

It's not clear that we have any non-circular way of picking out this aspect of mental states. Such attempts seem to always ground out in synonyms for "experience". So perhaps the best way to get clear on our target is to gesture at paradigmatic examples of phenomenal experience (e.g. pain, warmth, anger, visual experience) and draw a contrast with cases where experience seems to drop out of the picture (e.g. dreamless sleep, a coma).

We can also think about mental states in functional terms: in terms of the role they play in our cognitive economy, or in terms of the structural properties that explain their causal powers. Thinking in these terms is to think of the mind as a machine, whose operation is determined by its abstract causal structure.

The functional properties which feature in this mechanistic conception of the mind are, in principle, open to scientific investigation. In our own case, by carefully mapping perceptual inputs and behavioral outputs, by measuring patterns of activation in the brain, we can aim to construct a causal model of the human mind in terms of such functional features.

Such models typically invoke the functional notion of *access (a-) consciousness*[5]. When a mental state is a-conscious, its content is directly available for arbitrary use by central cognition, it's uniquely well-positioned to influence deliberate thought and behavior. Access consciousness works something like random-access memory (RAM) in a computer: its purpose is to keep information 'live', readily available for reference or further processing.

For the purpose of this paper, what matters most is that access consciousness is a functional concept. To say that a mental state is a-conscious is to say something about what it can *do*, about the sort of position it occupies in the mind's causal hierarchy. But I'll add two quick further clarifications.

Characterizations of access consciousness have sometimes been unnecessarily restrictive, by focusing on specific information-consuming capacities (e.g. speech production, abstract reasoning). Even organisms that lack complex linguistic or rational capacities can still surface information from cognitive subsystems, making that information centrally available for further processing. And a state can be *available* for reasoning or verbal report, without *actually* being used for those purposes. So access consciousness picks out the feature of cognitive systems that *underlies* or *explains* how a mental state can be used in reasoning or in the deliberate production of behavior (either verbal or non-verbal).

Characterizations of access consciousness can also be too permissive, by failing to distinguish between states that are *directly* vs. *indirectly* available. In a trivial sense, any of my beliefs are available to central cognition in and indirect sense. They need to be triggered or deliberately called up for them to 'come

[4]Block (1995)
[5]Block (1995)

online', to be available in a more direct sense. This distinction between directly available 'online' states and indirectly available 'offline' states can be spelled out in different ways, but, for now, just know that some such distinction needs to be made for there to be a meaningful distinction between a-conscious states and non-a-conscious states.[6]

So we have our two concepts of consciousness: "phenomenal consciousness" picks out the qualitative, how-it-feels sense of consciousness, and "access consciousness" picks out the functional, how-it-works sense of consciousness.

Before moving on, it's worth noting a couple of grammatical points about these terms.

First, phenomenal consciousness and access consciousness are categories we apply to mental states in virtue of their being conscious *at all*, without regard to their determinate character or specific content. To say that a mental state is p-conscious commits us to the claim that this state is associated with *some* determinate phenomenal character, but nothing more specific. This is why we can make sense of the claim that bats are p-conscious while being unable to conceive of the particular character of the corresponding p-conscious states. Similarly, to say that a state is a-conscious is not yet to commit it to playing any functional role other than the generic functional role picked out by "a-consciousness". For example, while you might think that a particular kind of mental state (e.g. pain) will necessarily involve a particular kind of functional role (e.g. the role which usually causes humans to shout "ouch!", the role which usually causes avoidance behavior, etc.), such particular roles are *not* implied by the attribution of a-consciousness. Instead, all a-conscious states (e.g. pains, tickles, memories, emotions) share the *very same* functional trait in virtue of being a-conscious: they are all directly available for use by central cognition.

Second, we typically take phenomenal consciousness and access consciousness to be *relational* properties, relating individual mental states to the subject to whom they belong. So when we say of a mental state that it's p-conscious/a-conscious, there should always be some subject for whom we're claiming that mental state is p-conscious/a-conscious.

2.2 The Conventional View

So we can think about the mind in phenomenal, how-it-feels terms or in functional, how-it-works terms.

How are these two ways of thinking about the mind related?

According to *conceptual dualism*, they aren't. Whereas phenomenal concepts are *first-personal* and *subjective*, functional concepts are *third-personal* and *objective*. As a result, we can't read off facts about phenomenology from facts about

[6] See Chalmers (1997) for a good discussion of these last two points.

functional architecture, and vice versa. While we may come to discover some pattern in the distribution of phenomenal properties and functional properties, no such pattern is revealed a priori, relying only on the concepts themselves.

Chalmers (1997) gives a nice, succinct expression of the view (as specifically applied to p- and a-consciousness):

> "Block's distinction between access consciousness and phenomenal consciousness (or experience) is very useful. There is clearly a conceptual distinction here, as illustrated by the facts that: (1) one can imagine access without experience and vice versa; (2) access can be observed straightforwardly, whereas experience cannot; and, most important, (3) access consciousness seems clearly amenable to cognitive explanation, whereas phenomenal consciousness is quite perplexing in this regard."

Conceptual dualism has played a particularly prominent role in shaping contemporary debates in philosophy of mind. This is due, in large part, to an influential series of arguments designed to demonstrate how functionalist theories of mind cannot explain phenomenal aspects of the mind.[7]

A commitment to conceptual dualism is now shared by philosophers who otherwise sharply disagree on issues of central importance. *Metaphysical* dualists of different varieties (e.g. epiphenomenalists, property dualists, ontological dualists) will, of course, find something attractive in conceptual dualism. But the view has also proven attractive to a posteriori physicalists, who think the link between consciousness and the functioning of physical systems must be established using standard empirical methods. Conceptual dualism even plays a key role in anti-realists' arguments against the existence of phenomenal consciousness: phenomenal properties don't make much sense on their own terms, or aren't compatible with our understanding of the natural world, and so should be abandoned.

To make conceptual dualism more precise, it'll help to say more about what it means for two concepts to be distinct, and how such a distinction can be established.

Concepts can differ in two ways. They can differ in their mode of presentation while still constraining each other's proper application (e.g. "is a triangle" and "is a shape with inner angles that sum to 180 degrees"). Or they can differ in a stronger sense, when the application of the one concept does not, a priori, constrain the application of the other (e.g. "is a square" and "is black", "is Clark Kent" and "is Superman"). In such cases, it's a priori conceivable for the

[7] Key works include Nagel (1974), Jackson (1982), Levine (1983), Block (1978, 1995), and Chalmers (1996).

properties picked out by each concept to vary independently.[8][9]

Conceptual dualists treat phenomenal concepts and functional concepts as distinct in this stronger sense. For instance, they'll claim that no amount of reflecting on the concepts of p-consciousness and a-consciousness will reveal any constraints on their actual overlap. When conceiving of a mental state as having phenomenal character, you can't simply reflect on this feature to discover any of its functional traits. And when conceiving of a mental state as a-conscious and thus playing a particular functional role, this feature alone can't help you determine whether that mental state has any phenomenal character.[10]

This can be spelled out in terms of missing *a priori entailments*:

> **No phenomenal-to-functional a priori entailments.** No description of a mental state in phenomenal terms a priori entails any description of that mental state in functional terms.
>
> **No functional-to-phenomenal a priori entailments.** No description of a mental state in functional terms a priori entails any description of that mental state in phenomenal terms.

Such claims about a priori entailments can be supported by conceivability tests: if X-but-not-Y is conceivable, then there's no X-to-Y a priori entailment; if X-but-not-Y is *in*conceivable, then there *is* an X-to-Y a priori entailment. And if either X-but-not-Y or Y-but-not-X is *in*conceivable (i.e. if there's either an X-to-Y or Y-to-X a priori entailment), then X and Y are *not* conceptually distinct.

When conceptual dualists apply these conceivability tests to p-consciousness and a-consciousness in particular, the result is:

> **That some mental state is a-conscious does not a priori entail that that mental state is also p-conscious.** This is shown by the conceivability of states that are a-conscious but not p-conscious. For instance, we can conceive of (philosophical) zombies that share our functional organization (and thus have states that are similarly *available* to central cogition) but lack any and all *experience*.

[8] Chalmers' treatment of conceivability (2002) provides a clear presentation of issues surrounding conceivability. In Chalmers hands, conceivability properly applies to *propositions*, or *statements*, which a reasoner either can or can't grasp. The relevant notion of conceivability, for this paper, is what Chalmers calls *ideal positive conceivability*, under which an ideal reasoner can conclusively conceive of the proposition in question.

[9] I won't, here, be engaging with disagreements over the nature of concepts, the a priori, and conceivability. Such disagreements shouldn't matter for the purpose of this paper. My aim is to bring out a problem that's *internal* to conceptual dualism, and I'm using these terms in a way that aligns with their usage in the relevant literature. Moreover, the challenge I raise can be posed using any theory under which limitations on what is thinkable correspond to conceptual limitations.

[10] Note that, in marking this *conceptual* distinction, we aren't hereby committed to a *metaphysical* distinction between p-consciousness and a-consciousness. Many accept this conceptual distinction while denying the corresponding metaphysical distinction. (Whether or not this view can be consistently maintained is a separate issue.)

> **That some mental state is p-conscious does not a priori entail that that mental state is also a-conscious.** This is shown by the conceivability of states that are p-conscious but not a-conscious. For instance, we can conceive of subjects who enjoy normal visual *experiences* which are not *directly available* to central cognition.

And so p-consciousness and a-consciousness, according to conceptual dualism, are conceptually distinct.

But this is wrong. In particular, the second conceivability test doesn't actually yield the verdict required by conceptual dualism.

While functional-to-phenomenal entailments have been subjected to a great amount of scrutiny, hardly any attention has been paid to phenomenal-to-functional entailments[11]. This is what I'll be focusing on for the rest of this paper.

But before getting on with it, I want to stave off a potential confusion: to conclude that a-consciousness is, a priori, *necessary* for p-consciousness is *not* to conclude anything about whether a-consciousness is *sufficient* for p-consciousness. That is, nothing in what I will go on to say should undermine or otherwise threaten the standard positive conceivability verdicts for absent qualia (and inverted qualia) cases. However, if p-consciousness without a-consciousness really is inconceivable, that *will* undermine any attempt to soley use absent qualia (or inverted qualia) cases to prematurely establish the sort of conceptual dualism outlined above. Put simply: one can accept that phenomenal consciousness is a functional concept without accepting that it is a *purely* functional concept.

3 Strange Experiences, Strange Subjects, and More

If conceptual dualism is right and there really aren't any a priori entailments from p-consciousness to a-consciousness, we should be able to cook up some

[11] A notable exception is Nagel's paper, "The Psychological Nexus", where he writes, "Although phenomenological features cannot be analysed behaviourally or functionally, their relation to their typical functional role in the production of behaviour is, in the outward direction, an a priori conceptual truth" (p. 451). Nagel is considering, in that paper, different strategies for shifting around and updating our phenomenal and functional concepts, such that any felt tension between them can be removed. Given this, it's not entirely clear whether he thinks any phenomenal-to-functional entailments can be read off our ordinary, everyday phenomenal concepts—but I do take that to be the suggestion.

Janet Levin has picked up on this suggestion in a series of papers (2002, 2007, 2008), but didn't endorse it herself, instead noting that, were the suggestion developed, it might provide a convenient way around the most pressing objections to a posteriori physicalism. But, to the best of my knowledge, Nagel's suggestion has otherwise remained unexplored.

truly strange scenarios that (according to conceptual dualism) will nevertheless be conceivable.

Let's return to the example I introduced at the beginning of the paper, the example of a *strange pain*. What makes this pain so strange is that although you *feel* the pain in all its glory (it's p-conscious), the pain is functionally isolated from your other conscious mental states and thus unavailable for any cognitive work (it's not a-conscious).

The strangeness, here, doesn't seem to hinge on the experience's being a *pain* rather than some other kind of experience with a different phenomenal character.[12] I can just as easily describe a *strange tickle* (i.e. an uncomfortable tickle that you *feel* on your left abdomen but is functionally isolated) or a *strange smell* (i.e. a pungent rotting odor that you *experience* but is functionally isolated).

Each of theses counts as a particular *strange experience*: a mental state that is p-conscious for some subject but also functionally isolated from, and therefore not a-conscious for, that subject. This gives us a general formula for taking any otherwise 'normal' p-conscious mental state and generating a corresponding strange experience by simply 'subtracting' the sorts of functional traits that would make it a-conscious.

Try it yourself: Take any aspect of you current experience. Think about what it would be like to continue enjoying that aspect of experience while subtracting all of the functional connections that put it into contact with central cognition, with the rest of your a-conscious mental states states.

We can try to 'estrange' progressively more and more of any given subject's phenomenal life until all of the subject's p-conscious mental states have been made strange. So, in addition to individual strange experiences, we can construct a series of progressively stranger cases:

> **Strange experience collection:** A subject has more than one strange experience at once. (e.g. strange pain + strange tickle + strange smell.)
>
> **Strange modality:** An entire sensory modality of a subject goes strange. (e.g. All visual experience is inaccessible.)
>
> **Strange subject:** Every p-conscious state of a subject goes strange. (i.e. the opposite of a philosophical zombie: all p-conscious states and no a-conscious states.)

[12] Nor does the strangeness hinge on the idiosyncratic functional profiles of particular conscious mental states. For instance, even if you think we can conceive of a pain which plays a radically different functional role while still conceiving of the pain *as* a pain (as David Lewis (1980) claims we can with his Mad Pain case), you should still be able to appreciate the strangeness in *my* pain example. Instead, I'll want to claim that you can't conceive of a pain as functionally isolated while conceiving of it as an *experience*. This all follows from something I noted about a-consciousness and p-consciousness earlier: they are general notions which can be applied without committing to the specific character or content of the mental state in question.

Since the strangest of these cases draws the sharpest strange/non-strange contrast, it's worth examining more closely.

What would a strange subject be like?

A strange subject is a sort of inversion of the classic zombie case.[13]. Whereas my zombie twin is a physical and functional duplicate that lacks p-consciousness, my strange twin will be a phenomenal duplicate that lacks a-consciousness.

Remember, your strange twin will, by definition, duplicate *all* aspects of your phenomenology. So if you were to enjoy the phenomenal life of your strange twin, things would feel just as they feel to you now. That is, your total p-conscious experience is subjectively indistinguishable from your strange twin's. Presuming that you know exactly what it is like to be you, you already know exactly what it is like to be your strange twin. It should therefore be fairly straightforward to conceive of *this* aspect of strange subjects.

If there's anything difficult about conceiving of our strange twins, that difficulty will stem from our attempt to reconcile this phenomenology with the erasure of a-consciousness. Again, strange subjects will share their phenomenology with us, but none of their p-conscious states are embedded in a cognitive system in the way ours are. To conceive of your strange twin is to conceive of subtracting a-consciousness from your own case—while leaving all of your phenomenology intact.

Can that really be done? I think not. At this point, we can identify a few different ways of worrying about the conceivability of strange subjects:

> **The Subjective Unity Worry.** If a strange subject has no access to any of its *own* p-conscious states, if its every experience is isolated from itself, it's difficult to project ourselves into the strange subject's point of view. Is there really any point of view left remaining if nothing is 'presented' to that point of view? Is there really any subject left remaining if the various experiences that 'come together' in a subject can't bear any functional relations to one another and to the subject who owns them? To have a subject at all, various experiences may have to be functionally integrated in order to form a single cohesive subject. In this case, the various aspects of a subject's overall experience might be said to bear unity relations to each other. And it may also be said that the subject bears an ownership

[13]Though strange subjects shouldn't be confused with the "anti-zombies" (Frankish, 2007) or "reverse-zombies" (Brown, 2010) discussed elsewhere. Interestingly, Frankish and Brown had something very different in mind in their discussion of opposite-of-zombie cases. Frankish's anti-zombies live in a world where "consciousness is a physical phenomenon, supervening metaphysically on the world's microphysical features" (p. 653). And Brown discusses two different 'reverse-zombie' cases: 'zoombies', "creatures which are nonphysically identical to me in every respect and lack any nonphysical phenomenal consciousness"; and 'shombies', "creature[s] that [are] micro-physically identical to me, [have] conscious experience, and [are] completely physical" (pp. 50-51). Strange twins are phenomenally identical to you, but they lack the functional structure which, in you, accompanies conscious experience.

relation to its experiences. And it's precisely these features that seem threatened once a-consciousness is removed.

The Mental Action Worry. Even when thinking of a purely phenomenal mind (e.g. something like a Cartesian Ego), we think about such minds as capable of mental action. Minds can *do* things: they make decisions; they form judgments; they rehearse mental routines—and all this can be understood entirely from the subject's own internal perspective. It's not clear that *any* of this would be possible for a subject that lacked any and all functional organization. Could a strange subject ever perform a calculation or pursue some action, mental or otherwise? Could a strange subject *feel* like there's some mental activity going on without there actually being any corresponding mental activity? In our own case, it at least *seems* like such mental activities make a contribution to our overall experience. Could this sort of experience ever be illusory? And if so, are we only ever experiencing the effects of 'mental action' without performing those mental actions ourselves? If there is a difference between only *seeming* to perform mental actions and *actually* performing mental actions, how could we ever tell the difference?

The Temporal Relations Worry. We typically conceive of subjects as enjoying a temporally-extended life. Moreover, we typically think that the experiences of a subject at one time are causally related to the experiences of the same subject at a later time. So subjects, as we typically conceive of them, seem to be temporally structured. Could a strange subject be temporally structured in this way? When a strange subject feels a painful sensation (e.g. the sensation of a punch to the gut), then has a visual experience as of the cause of that sensation (e.g. seeing (or seeming to see) their friend Frank punching them), then has an emotional 'response' (e.g. anger at Frank), then feels a desire to take a particular action (e.g. a desire to punch Frank back), and then experiences the performance of that action (e.g the punching of Frank)—could any of these experiences be causally related (given the constraint on functional organization)? If not, can we make sense of a single strange subject having a sequence of experiences like this? Or must strange subjects, as a result of lacking the functional structure which would maintain them through time, exist only instantaneously?

I, myself, don't feel equally drawn in by each worry. (I feel the pull of the subjective unity worry much more than the other two.) But any of these worries can provoke the intuition that conceiving of a phenomenal mind is to conceive of an underlying functional structure in which various experiences are embedded and through which they bear relations to one another. We may notice that we don't think of the ebb and flow of experience as separable from structural or causal relations, that we often *do* apply a functional lens in thinking about

the collection of experiences 'housed' in a given subject, and that this way of thinking isn't optional, but, rather, already implicated in the basic conception of experiences and experiencers. To endorse this line of thinking is to do endorse what I'll call *the structured subject intuition*. Perhaps there are other ways of questioning the conceivability of strange experience cases, but what's distinctive about this kind of approach is its focus on the conceptual priority of subjects, as systems with constraints on their functional organization.

I focus on the structured subjection intuition because I think it provides a compelling explanation of *why* strange experiences are inconceivable. But *any* way of questioning the conceivability of strange experience cases will be a way of questioning the conceptual independence of phenomenal consciousness from anything functional. To take strange experience cases seriously is to take seriously the idea that thinking in phenomenal terms necessarily involves thinking in functional terms, that judgments about phenomenology constrain judgments about functional structure, that there are a priori entailments from phenomenal descriptions to functional descriptions. Or, to put it another way, if strange experience cases are inconceivable, that means that we can learn something about the functional structure of a subject just by learning about her phenomenology.

A few additional features of the structured subject intuition are worth noting.

First, to think of a subject as structured or as admitting of functional organization is not yet to think of that subject as having *physical* structure. Standardly, facts about functional structure and organization are understood to capture the abstract causal relations that hold between the different parts of a system. To rehearse a familiar point about the multiple realizability of functional properties, two systems can have the same functional properties (at least at a certain level of description) but do so with parts that are, physically, quite different from one another. And if purely phenomenal systems are conceivable, a phenomenal system could, in principle, have the same functional properties as some physical system. Seeing that this is a conceptual possibility can help clarify that the structured subject intuition takes no sides in the debate between physicalists and dualists.

Second, while I've used the example of strange subjects to bring out the structured subject intuition, it can also be brought out for any case lying on the spectrum of strangeness introduced above. The difficulties we run into with strange subjects should be precisely the same as those we have with single strange experiences. Each worry at the global level should turn up at the local level:

> **The Local Subjective Unity Worry.** How can a strange experience (like a particular strange pain) be part of or be owned by the subject for whom it is an experience?
>
> **The Local Mental Action Worry.** How could a strange experience play any role in mental action?
>
> **The Local Temporal Relations Worry:** How could a strange

> experience be part of any temporally extended causal chain or endure through time?

But any underlying uneasiness is best brought out in the case of strange subjects. Whereas subjects are isolated from only *one* of their mental states when a particular experience of theirs is strange, strange subjects are isolated from the *entirety* of their phenomenology. It's one thing to be estranged from a small portion of one's total experience. It's something else to be estranged from the *entirety* of one's experience. And while we may find clever ways to convince ourselves to allow 'hidden' phenomenology in particular instances, it's harder to allow *all* of our phenomenology to be 'hidden'. And that global-level clarity should, consequently, make us reconsider any stories we tell to paper over local-level issues.

Last, while my discussion will be focused on the broad class of strange experiences described above (ranging from individual strange experiences to strange subjects), the structured subject intuition can *also* be brought out by other related thought experiments. Consider:

> **Strange Modality Expansion:** A subject has an additional strange modality of which they are not aware. Consider, for instance, that you, yourself, have a strange modality which gives you a bat's sonar experiences. But these bat-like experiences would, nonetheless, be inaccessible to you. (And if you don't think this is your actual situation, how do you know that this is not your actual situation?)
>
> **Strange Swapping:** Two different subjects are a-conscious of *each other's* p-conscious states. (Or vice versa: two different subjects are p-conscious of *each other's* a-conscious states.) For instance, consider a case in which 'you' are currently experiencing the life of someone else while also taking actions informed by 'your own' mental states and making decisions about your own life.
>
> **Strange God:** Consider a subject who has every experience in the world (i.e. my experiences, your experiences, a bat's experiences, a dolphin's experiences, etc.) but is unable to access any of them. Similarly, consider a subject who has all experiences in the world but can only access the experiences of a single organism (e.g. you experience everything but are only a-conscious of 'your' p-conscious states.)

Insofar as any strange experience case makes you pause and question whether it actually makes sense, it's worth figuring out whether or not that means you are implicitly committed to some a priori functional constraint on experience. With the structured subject intuition, I tried give and initial sketch of a positive alternative to conceptual dualism. That alternative can take different forms, depending on how exactly you draw the distinction between conceivable cases and inconceivable strange experience cases. (These different ways of drawing the distinction will map onto different conceptions of a-consciousness, or different

functional traits that are entailed by p-consciousness.) But the particulars of that positive account need not be settled in order to use strange experience to make a negative case against conceptual dualism.

4 Can Strange Experiences be Made Palatable?

Strange experience cases present conceptual dualists with a challenge. Under conceptual dualism, phenomenal descriptions can't a priori entail functional descriptions, and vice versa. But strange experience cases seem to show that phenomenally conscious states must, a priori, satisfy certain functional constraints. So it seems like the how-it-feels thinking isn't actually separable from how-it-works thinking. And so conceptual dualism seems false. To block this challenge, conceptual dualists need to explain why it's a mistake to think strange experiences inconceivable.

So how might we respond to the strange experience challenge on behalf of conceptual dualism?

4.1 Radical A Priori Skepticism

We can pretty quickly rule out strategies which try to undermine the challenge by way of 'radical' skepticism about conceivability and the a priori:

"Why be so sure that there's some fact of the matter about whether strange subjects are conceivable? Not all concepts are well-behaved: they sometimes break down, no longer guide us, and leave us unsure about how to apply them. It would be a mistake to insist that there be some clear verdict on whether a concept applies in such problem cases."

This approach can't be used in the present context, in support of conceptual dualism. It proves too much: it undermines the theoretical framework used to articulate the central conceptual dualist commitment. Conceptual dualism *needs* positive conceivability verdicts for any arrangement of phenomenal and functional properties, otherwise we lose the positive justification for being conceptual dualists.

4.2 Different Notions of Access

We might question whether strange experiences, as described, *really* present us with a case of p-consciousness without a-consciousness. If we make available a revised and updated conception of a-consciousness—one which makes clearer the stark contrast between a-consciousness and p-consciousness—perhaps it will be easier to see that strange experiences really are conceivable. Here are two candidates for a clarified notion of a-consciousness:

Accessed/Accessible

"We can distinguish between mental states that are *actually accessed* from those which are *merely accessible*. And if we understand 'a-conscious' to mean 'actually accessed', then it's easier to accept that there can be p-conscious states that don't become a-conscious."

Block, for instance, pursues something like this strategy:

> "Importantly, the [phenomenal] overflow argument does not claim that any of the items in the array are cognitively inaccessible, but rather that necessarily most are unaccessed. For comparison, consider the following: not everyone can win the lottery; however, this does not show that for any particular contestant the lottery is unwinnable. In other words, to say that necessarily most items in the array are not accessed is not to say that any are inaccessible." (Block 2011, p. 576)

> "Many critics of the overflow argument seem to think that a vote for overflow is a vote for inaccessible consciousness. ... However, as pointed out earlier, the fact that necessarily most items are not accessed does not entail inaccessibility of any items." (Block 2011, p. 583)

Fancy/Folky

"We can distinguish between *fancy* 'theory-laden' formulations of a-consciousness that rest on work in cognitive science and *folky* formulations of a-consciousness that capture some pre-theoretical, everyday notion of 'access'. Since the theory-laden way of conceiving of mental states may be far removed from how the folk think about them (it'll involve lots of additional commitments about the structure of our mental machinery), this sense of a-consciousness clearly isn't tied up with our everyday p-consciousness concept. And so its easier to see how the may dissociate."

Any such attempt to shift our notion of a-consciousness will have the same basic problem: it merely relocates, rather than removes, the present challenge. I'll specifically address the accessed/accessible version of this objection, but my reply to both versions is essentially the same.

Replacing a constraint that *all p-conscious states must be accessed* with a constraint that *all p-conscious states must be accessible* does nothing to remove the worry that p-conscious states might have *some* a priori functional constraint. Accessibility is, after all, a functional notion. You could, of course, drop the accessibility constraint too, but then the accessed/accessible distinction isn't doing any work here to make it any *easier* to address the challenge. So I suspect that those drawn to this objection are not appreciating the nature of the strange experience challenge.

4.3 Dimmed Experience

I suspect that some people, when confronted with my strange experience cases, will try to point to a class of *actual* cases of p-consciousness without a-consciousness with which we're already well-acquainted. For instance, Block has also made appeals along these lines:

> "Suppose you are engaged in intense conversation when suddenly at noon you realize that right outside your window there is—and has been for some time—a deafening pneumatic drill digging up the street. You were aware of the noise all along, but only at noon are you consciously aware of it. That is, you were P-conscious of the noise all along, but at noon you are both P-conscious and A-conscious of it. . . . [This] is a pure case of P-consciousness without A-consciousness" (Block 1995, p. 234)

It's easy to find further examples of this sort of thing, instances where some experience has receded into the periphery of our phenomenology. We typically have no difficulty calling such 'dimmed' experiences *experiences*, but such experiences may not be a-conscious in the way that other experiences are. As the thought goes, since such experiences are not salient enough, they may easily pass by without us actually 'accessing' them. So we, perhaps, have *actual* cases of p-consciousness without a-consciousness—even better than merely conceivable cases.

Dimmed experiences, intuitively, are available in ways that completely unconscious states and processes are not available. We think, "there was some aspect of my experience that has been there for awhile and nothing was stopping me from foregrounding that experience". Unconscious states and processes, we think, are unconscious precisely because they aren't 'at the ready' in this way. And so there's a functional difference between merely-dimmed experiences and fully-extinguished experiences.

So this objection fails for the same basic reason the last objection failed. Dimmed experiences, as described, look more like 'unaccessed' but still 'accessible' mental states. And I think it would be much less clear that these were really experiences if they were totally unavailable/inaccessible. So it doesn't address the claim that p-consciousness is tied to the functional notion of accessibility.

4.4 Unimaginable, Not Inconceivable

Leaving disagreements over the notion of a-consciousness behind, conceptual dualists might pursue a line of defense centering on the notion of conceivability:

"Are we sure that we run into any tension just in *conceiving* of an inaccessible experience? It's more likely that what gets us into trouble is an act of *imagination*—and that's separable from any act of *conceiving*. Imagining an experience

involves simulating, on our own hardware, what that experience feels like. When we try to simulate what a strange experience would feel like and then 'check on' this simulated experience, we, unsurprisingly, *always* find such experiences to be a-conscious—'checking on' an experience is a paradigmatically a-conscious action! That we find ourselves lapsing into imagination in these cases is understandable. It's quite a handy tool when trying to understand others' experiences. But in the present context, this knee-jerk reliance on imagination leads us astray, and we're mistakenly pulled towards the structured subject intuition as a result."

Before I get to my main response, I want to note that this line of reasoning doesn't yet provide any support for a *positive* conceivability verdict. It merely explains away the supposed *negative* conceivability verdict. So for this objection to work, it would need an additional, independently motivated story in support of the positive conceivability claim. Without such a story, we've lost a key motivation for adopting conceptual dualism over alternative views.

But to address the substance of this objection: it's important to recognize that our concept of p-consciousness can be applied without actively imagining or conceiving of any *particular* kind of experience. This follows from a general point about p-consciousness briefly gestured at earlier: p-consciousness is a determinable property, not a determinate property. Consequently, it's easy to conceive of a state being p-conscious while leaving its precise phenomenal character unspecified. For instance, we can make sense of the claim that bats enjoy uniquely batty experiences, and this doesn't require any insight into the particular nature of any of these batty experiences. Furthermore, I take it that bat-owned strange experiences are just as inconceivable as human-owned strange experiences. Could a bat enjoy an experience associated with some functionally isolated state, like some state in a different bat's head? Such a case is just as implausible as the strange pain case described earlier. In this way, I think we can demonstrate that the imaginability/unimaginability of particular kinds of experiences plays no essential role in our assessment of strange experience cases.

4.5 Constraints on Rationality, Not Phenomenality

A conceptual dualist might want to grant that there's something difficult about conceiving of strange experience cases, but locate the source of that difficulty elsewhere:

"What's strange about these cases is not that they are *incoherent* but that they ask us to take up the point of view of a subject who suffers from an unfamiliar, severe form of irrationality. Strange subjects fail, in systematic ways, to draw the correct inferences from experience, or otherwise marshal the contents of their mind in service of cognitive tasks. So maybe we can't conceive of a subject who undergoes a strange experience while *also* conforming to certain rational norms. But so long as we allow that such subjects may be irrational in unfamiliar ways, strange experience cases should escape any conceivability challenges."

Under this objection, there's still something to be learned from strange experience cases: they demonstrate a conceptual connection between rationality and a-consciousness. But, importantly, this doesn't undermine conceptual dualism.

I don't think this objection does much to avoid or account for the seeming difficulty of ascribing strange experiences to the subjects for whom they are inaccessible.[14] We can attribute irrationality only after we deem these experiences to belong to the subject, and I don't think we can do that.

But even so, charges of irrationality are misplaced here.

You are rationally criticizable insofar as you fail to integrate all the information available to you in a systematic, consistent way. But in a strange experience case, the subject does *not* have access to key information, and so she has not even been given a chance respond to and act on this information in a rational way. It's hard to find anything 'irrational' about not noticing that you're having experiences that are in-principle-unnoticeable.

Consider an analogy: A detective that has every clue, but fails to put them together, has made a regrettable rational error. But a detective who failed to collect all the necessary evidence in the first place hasn't made a *rational* error. Perhaps the detective can be faulted for failing to retrieve evidence which was, in fact, retrievable. But, if we're to follow the analogy through, strange experiences are akin to irretrievable evidence. And I can't see how we could fault a detective for failing to recover some piece of evidence that was, for instance, destroyed before she arrived at the crime scene.

Perhaps this objection relies on an externalist conception of rationality, a conception under which you can be faulted for failing to know about any fact out in the world that you 'should' know about, even for facts that you couldn't possibly know about. But then irrationality is cheap. And if irrationality is cheap, it's hard to see how irrationality can be the source of a kind of tension that is unique to strange experience cases.

4.6 Constraints on Phenomenal Subjects, Not Phenomenality

At this point, conceptual dualists might complain that strange experience cases weren't really set up to test their central commitment:

"Strange experience cases may reveal a priori constraints on phenomenal subjecthood, but such constraints shouldn't be confused for constraints on phenomenal experience itself. That is, strange experiences don't show that the phenomenal, as a conceptual domain, must be understood in functional terms; rather, they just show that some particular functional arrangement of phenomenal experiences implicates functional facts. But that's hardly interesting. And it doesn't

[14] And so this objection may just collapse into the 'bare denial' strategy, outlined below.

directly challenge the idea that any arrangement of the atomic phenomenal facts is consistent with any set of functional facts."

I have two replies. First, whether some set of phenomenal experiences is unified together in, or belongs to some phenomenal subject should count as a phenomenal fact. That is, *that S is an experiencer* seems to fall clearly on the phenomenal side of any phenomenal/functional line. And all that is required to undermine conceptual dualism is for there to be *some* phenomenal facts which entail functional facts. So even if "S is a subject of phenomenal experience" a priori entails "S is a-conscious", I'll have successfully undermined conceptual dualism. And once we break the seal and allow for *some* a priori connections of this kind, we may rightly question the motivations for keeping these conceptual domains separate in the first place.

Second, it is not clear that we can conceive of individual experiences that do not belong to a subject. This is a familiar point, and it has been made a number of times, in different ways. But to put it simply, many find it plausible that: if x is a p-conscious mental state, there must exist some subject S which is p-conscious of x. If this is right, then a functional constraint on phenomenal subjecthood would, consequently, constrain the entire phenomenal domain.

But is it right? I'm not sure there's much additional I can say here, either for or against. If phenomenal consciousness is a primitive concept, we should be able to immediately read this off of our basic concept of experience. To have the concept of experience just is to be able to recognize such truths. It should be as obvious that p-conscious states must belong to phenomenal subjects as it is obvious that beliefs and desires must belong to rational agents. The only way to test this would be consider cases not unlike those we've already been considering: cases of subject-less experiences. And so it's not clear that this objection actually represents a substantive argumentative move.

4.7 Bare Denial and Beyond

For conceptual dualists who remain unmoved by the strange experience challenge, one last refuge remains:

"You're right that the above objections can't be used to decisively cut off skepticism about conceptual dualism. But I still don't see what's so unacceptable about embracing the conceivability of strange experiences. That's just how these things go sometimes. When intuitions diverge over basic conceptual starting points—presuming that we really have the same concept in mind and are not talking past one another—all that's left is a familiar sort of bare disagreement. And while it doesn't look like this issue can be settled by any further analysis, that doesn't mean we don't already have enough evidence to support a positive conceivability verdict. You say you can't conceive of inaccessible experiences, but you're *just wrong*. And recognizing this involves nothing more than exercising one's competence in wielding the concepts involved."

Indeed, this is how these things go sometimes. And I agree that we should embrace the basicness of this issue. (If I were a conceptual dualist, I'd go with this non-objection objection.) But I want to note: a stalemate, here, would be significant. Conceptual dualism is frequently taken as an uncontroversial background commitment in philosophical discussions of consciousness. By reexamining conceptual dualism and dislodging it from its privileged position—by foregrounding an issue that had been obscured and ignored—we open up and make available views that weren't previously given proper consideration. A reconsideration of such views can have a number of downstream consequences for issues the framing of which took conceptual dualism for granted, either implicitly or explicitly. And by reshaping the terrain of the debate in this way, newly-available views may consequently be seen as more attractive—they may help solve problems which, under the old framework, were thought to be unsolvable or otherwise very difficult. Even small victories of this sort would provide abductive support for rejecting conceptual dualism.

With that meta-theoretical story out of the way, I'll wrap things up with a brief sketch out the alternative path I'm proposing we take.

5 Following the Structured Subject Intuition

If strange experiences are deemed inconceivable, that suggests that p-conscious mental states must, a priori, be a-conscious. Because the notion of a-consciousness was, itself, left a little uncertain, the precise nature of the functional constraints on experience still need to be spelled out. (As I see it, we can develop a more precise notion of a-consciousness by starting from strange experience scenarios and seeing which functional features need to be added back in to make these scenarios conceivable again.) However, without get hung up on this leftover task, we can still say something about the basic idea that there are functional constraints built into the notion of experience.

A few clarifications should allow us to better appreciate the basic view I have in mind.

First, this view should not be mistaken for analytical functionalism, which is a *much* stronger view. While analytical functionalism holds that *all* mental thought and talk can be *reduced* to functional though and talk, the view in question only claims that the determinable concept of "experience"/"phenomenal consciousness" brings along with it certain functional constraints. It need not be the case that such a concept can be given an exhaustive analysis in functional terms.[15] Nor does the view require that more determinate phenomenal concepts (e.g. "pain", "visual experience", "joy") bring along with them more determinate functional constraints. So while this view will say that a token pain state is, a

[15] I, myself, would reject the possibility of such an exhaustive analysis, but this possibility is left open by the view.

priori, a-conscious, it does not tell us anything about the functional profile that distinguishes pain experience from other kinds of experience.

Second, as I stressed earlier, this view is only concerned with one of the two directions of a priori entailment denied by conceptual dualists: the phenomenal-to-functional direction. Accepting an entailment of this kind does not commit us to accepting any entailments in the other direction. So, for instance, this view doesn't rule out the conceivability of philosophical zombies.[16]

Third, it's important to remember that this is a view about the *concept* of—not the metaphysics of—experience. So the view is committed neither to materialism (since one can, in principle, combine functionalism with dualism), nor to phenomenal realism (since it remains to be seen if there's anything in reality to which our concept of experience properly applies).

That said, a revision to our understanding of the concept of experience will necessarily trigger revisions to the metaphysics and epistemology of experience. In particular, this view introduces, for free, functional 'hooks' for experience, which can simplify otherwise-frustrating methodological hurdles that frequently stymie the science and metaphysics of mind. For example, we might immediately rule out panpsychism as a plausible metaphysical view, since many of the states which would count as p-conscious under this view violate the sorts of functional constraints being introduced. We might also rule out views like Dan Dennett's multiple drafts theory of consciousness, since such a theory would, at best, require positing multiple minds for which each 'draft' stands in the right functional relation.

How exactly to fill in the details of this view is a project for another time. But my hope is that, with this paper, I've raised doubts about the plausibility of conceptual dualism and opened the door for an attractive alternative.

6 References

Block, Ned (1978). Troubles with functionalism. Minnesota Studies in the Philosophy of Science 9:261-325.

Block, Ned (1995). On a confusion about a function of consciousness. Brain and Behavioral Sciences 18 (2):227—247.

Block, Ned (2011). Perceptual consciousness overflows cognitive access. Trends in Cognitive Sciences 15 (12):567-575.

Brown, Richard (2010). Deprioritizing the A Priori Arguments against Physicalism. Journal of Consciousness Studies 17 (3-4):47-69.

[16] I, myself, would reject any functional-to-phenomenal entailments, but the view I'm laying our requires no such commitment.

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

Chalmers, David J. (1997). Availability: The cognitive basis of experience? In Ned Block, Owen J. Flanagan & Guven Guzeldere (eds.), The Nature of Consciousness. MIT Press 148-149.

Chalmers, David J. (2002). Does conceivability entail possibility? In Tamar S. Gendler & John Hawthorne (eds.), Conceivability and Possibility. Oxford University Press 145-200.

Frankish, Keith (2007). The anti-zombie argument. Philosophical Quarterly 57 (229):650–666.

Jackson, Frank (1982). Epiphenomenal qualia. Philosophical Quarterly 32 (April):127-136.

Levin, Janet (2002). Is conceptual analysis needed for the reduction of qualitative states? *Philosophy and Phenomenological Research* 64 (3):571-591.

Levin, Janet (2007). Nagel vs. Nagel on the nature of phenomenal concepts. *Ratio* 20 (3):293–307.

Levin, Janet (2008). Taking type-b materialism seriously. *Mind and Language* 23 (4):402-425.

Lewis, David (1980). Mad pain and Martian pain. In Ned Block (ed.), Readings in the Philosophy of Psychology. Harvard University Press 216-222.

Levine, Joseph (1983). Materialism and qualia: The explanatory gap. Pacific Philosophical Quarterly 64 (October):354-61.

Nagel, Thomas (1974). What is it like to be a bat? Philosophical Review 83 (October):435-50.

Nagel, T. (2000). The Psychophysical Nexus. In Paul Boghossian and Christopher Peacocke (Ed.), New Essays on the A Priori. Oxford University Press 434-72.

Phillips, Ian (2018). The methodological puzzle of phenomenal consciousness. Philosophical Transactions of the Royal Society of London B. (373): 20170347.