Strange Experience: Why Experience Without Access Makes No Sense

Jackson Kernion

August 24, 2018

Abstract

According to a widely held view, we have two separate ways of thinking about consciousness. We can think of 'conscious' mental states in phenomenal terms–as feeling a certain way or as having certain qualitative aspects–or we can think of 'conscious' mental states in *functional* terms–as being available for arbitrary use. Importantly, these are *conceptually distinct* in the sense that we can't simply reflect on the concepts involved to learn about how they do/don't actually overlap. But I think we have good reason to reject this view: we cannot conceive of an experience that is functionally isolated from–and therefore unavailable to–the subject for whom it is an experience. If this is right, it looks like the phenomenal and functional senses of consciousness aren't totally distinct after all, since it suggests that there are a priori functional constraints on experience.

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 neural architecture just by learning about whether/what that organism feels, and vice versa.

But this is a mistake–a mistake which has generated 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),

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

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 this 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*.

Of course, this isolated pain won't seem so strange were we to conceive of it 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 within your purview, as something that you're experiencing) and in another sense *not yours* (as something lying outside your purview, as something functionally isolated and 'removed' from your other mental states).

Perhaps my description of this case is illegitimate, or involves some confusion that ought to be corrected. But I think sustained reflection will reveal that the central tension of this scenario cannot be resolved. This would be instructive: it would suggest that a functional constraint must be built into our 'experience' concept, that having a pain–or *any* kind of experience–necessarily involves having a certain kind of 'access' to it. Without meeting such a constraint, we can't make sense of experiences *as* experiences.

This is the kind of view I want to make available with this paper. But such a view stands opposed to the conventional wisdom that 'experience' and 'access' are fundamentally different concepts. This unconventional view would, consequently, obviate recent debates about whether phenomenology overflows cognitive access[2] and shift some of the key theoretical suppositions that inform philosophical and scientific investigations into the nature of consciousness. I'll leave the discussion of these downstream consequences for another time. My focus in this paper will be the plausibility of the standard view mentioned at the outset of this paper. Ultimately, I think reflecting on cases of strange experience will make such a view untenable.

The plan for the rest of the paper: In section 2, I lay out more precisely the conventional view I'm targeting. In section 3, I first generalize and expand the strange pain case into a more robust skeptical challenge, and then provide some

[2]Block (2011) kicked things off. Phillips (2018) offers a useful overview how the debate has developed, and how the issue has remained unresolved. [ADD]

reasons for thinking the challenge insurmountable. In section 4, I go over some strategies for pushing back against this challenge, and say why none of them work. In a brief final section, I say a little more about the sort of view we should adopt in light of the standard 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 mental states.

First, we can think of mental states in phenomenal terms: in terms of *what it feels like* to be in that state, in terms of its *subjective qualitative character*. To think of a mental state in these terms is to think of it as *phenomenally* (p-) *conscious*.

It's not clear that we have any non-circular way of picking out this aspect of the mind. 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 the causal connections linking its component parts.

The functional properties which feature in this mechanistic conception of mind are, in principle, open to scientific study. 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 functional model of the human mind.

In such models, one functional property is particularly salient: *access* (a-) *consciousness*. When a mental state is a-conscious, its content is made available for arbitrary use by central cognition, its uniquely well-positioned to directly influence deliberate thought and behavior. Characterizations of a-consciousness have sometimes focused on specific information-consuming capacities (e.g. speech production, abstract rational thought), but this characterization is too restrictive. Even organisms who lack language/complex rational thought can still surface information from subsystems, making that information globally available for further general-purpose processing. And so a-consciousness picks out the feature of cognitive systems that *underlies* or *explains* how a mental state can be made

available for use in reasoning or in the deliberate production of behavior (either verbal or non-verbal).[3]

(This characterization of a-consciousness is fairly crude, but that's in line with how it is standardly characterized. And we shouldn't need anything more precise for the line of reasoning I want to pursue in this paper, which will draw primarily on the phenomenal/functional contrast.)

It's worth noting a couple grammatical points about p-consciousness and a-consciousness.

First, p-consciousness and a-consciousness are categories we apply to mental states in virtue of them being 'conscious' *at all*. To say that a state is p-conscious commits us to the claim that this mental 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 make sense 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 more specific than the general functional role picked out by 'a-consciousness'. For example, while you might think that certain kinds of mental states (e.g. pain) will necessarily involve certain kinds of functional roles (e.g. the role which usually causes humans to shout "ouch!", or the role which usually causes avoidance behavior), such 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 p-consciousness/a-consciousness to be *relational* properties, relating individual mental states to the subject/system 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/system for which we're claiming it's p-conscious/a-conscious.[4]

2.2 Conceptual Dualism: First pass

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 one prominent view, they aren't! The concepts used in conceiving of the mind in phenomenal terms are distinct from the concepts used to conceive of the mind in functional terms. Whereas phenomenal concepts are *first-personal* and *subjective*, functional concepts are *third-personal*, and *objective*. Phenomenal concepts can be used to capture the *essential, intrinsic nature* of mental states,

[3] See Chalmers (1997) for further discussion of these points.

[4] Some might object to this relational interpretation, but I don't think much hinges on this–I just want to clarify how I'll be using these terms.

while functional concepts can be used to capture *inessential, extrinsic relations*. As a result, we can't read off facts about phenomenology from facts about functional architecture, and vice versa. This is what I'll be calling *conceptual dualism*.[5]

A commitment to conceptual dualism is shared by philosophers who otherwise disagree sharply on issues of central importance. Conceptual dualism will, of course, be attractive to ontological dualists, but it's also attractive to a posteriori physicalists, who think the physical basis of consciousness must be discovered empirically. Conceptual dualism also plays a key role in arguments for phenomenal anti-realism: 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.

That should give you a flavor of the view I want to target. But to make the view more precise, it'll help to take a closer look at what it means for two concepts to be distinct, and how such a distinction can be established.

2.3 Conceptual Distinctness

Claims about conceptual distinctness can be easily misunderstood, since they can be taken in two different ways.

First, it can be claimed that two concepts are *weakly distinct* when they differ in their mode of presentation. For example, the concepts 'square' and 'regular diamond' are weakly distinct because the 'aspects' by which each concept picks out its referent are different.[6]

But, of course, to conceive of a square will be to conceive of something to which the concept 'regular diamond' properly applies. And so some may want to say that the two concepts are 'not *really* distinct': though their conceptual overlap may not be *psychologically* immediate, it is *logically* immediate. Someone wielding both the 'square' and 'regular diamond' concepts should be able to discover their conceptual overlap merely through a certain kind of reflection, using the very same mental or cognitive resources required to grasp those concepts in the first place. We can capture this idea is in terms of *a priori entailments*. Because "this is a square" a priori entails "this is a regular diamond" (and vice versa), the distinction between 'square' and 'regular diamond' is weaker than other conceptual distinctions.

But when such a priori entailments *are* ruled out, the two concepts involved will be 'distinct' in a *stronger* sense: the two concepts may, in fact, overlap in

[5] I have in mind, here, the influential views of Nagel (1974), Block (1995), and Chalmers (1996) in particular.

[6] Another example of (merely) weakly distinct concepts: 'triangle'/'shape with inner angles that sum to 180 degrees'. Note that, as I've laid out this distinction, all strongly distinct concepts will also be weakly distinct.

their extension, but their overlap is not guaranteed a priori.[7] So 'square' and 'regular diamond' aren't conceptually distinct in this stronger sense, but the concepts 'square' and 'black' are. You can't conceive of a square that's not a regular diamond. But you *can* conceive of an square that's not black, or an object that's black but not a square. The former case (a non-diamond square) involves a kind of inconsistency that the latter cases do not. Since there's no a priori connection between an object being black and an object being a square, the proper application of the one concept does not, in principle, constrain the proper application of the other.

2.4 Conceptual Dualism: Second Pass

Conceptual dualists treats phenomenal properties and functional properties as conceptually distinct in this stronger sense. (Going forward, I'll only be talking about this stronger sense, dropping explicit 'in the stronger sense' qualifiers.) For instance, they'll claim that no amount of reflecting on the concepts 'p-consciousness' and 'a-consciousness' will reveal heretofore unknown constraints on their actual or potential overlap. When conceiving of a mental state as having phenomenal character, you can't discover (by way of reflecting on this feature alone) any functional properties that would fix if/how such a state would be directly available for use. 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.[8]

Conceivability plays a crucial justificatory role here. We track the a priori connections that do/don't hold between concepts by applying a simple conceivability test: if X-but-not-Y and Y-but-not-X are conceivable, then X and Y are conceptually distinct. And if either X-but-not-Y or Y-but-not-X is *in*conceivable, X and Y are *not* conceptually distinct.[9]

So to map this out: the conceptual distinctness of phenomenal properties and functional properties is to be understood in terms of missing a priori entailments. And the absence/presence of a priori entailments are shown by way of conceivability tests. This forms the theoretical backdrop of the view I want to assess in this paper.[10]

[7] Further examples of strongly distinct concepts: 'located in region X'/'has shape Y', 'Superman'/'Clark Kent', 'Water'/'H2O'.

[8] 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.)

[9] 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.

[10] I won't, here, be engaging with disagreements over the nature of concepts, the a priori, and conceivability. Such disagreements shouldn't much matter for our purposes. My aim is to

When conceptual dualists apply this conceivability test to p- and a-consciousness, they find:

That some mental state is a-conscious does not a priori entail that 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 *available* for use in the same way) but lack any and all *experience*.

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* for general cognitive tasks.

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

But this is wrong. In particular, I don't think the second conceivability test yields the verdict that is consistent with conceptual dualism. While functional-to-phenomenal entailments have been subjected to a great amount of scrutiny, hardly any attention, in my estimation, has been paid to phenomenal-to-functional entailments. And that's what I'll be focusing on for the rest of this paper.

3 Strange Experiences, Strange Subjects, and More

If conceptual dualism is right and there really aren't any a priori entailments between p-consciousness and a-consciousness, we should be able to cook up some truly strange scenarios that (according to conceptual dualism) will nevertheless be conceivable.

Let's return to the example I introduced earlier, an example of what I'll call *strange pain*. What makes this pain so strange is that although you *feel* the pain in all it's glory (i.e. it's p-conscious), the pain is functionally isolated from your other conscious mental states and thus unavailable to you (i.e. it's not a-conscious).

Strange pain may be particularly evocative, but its strangeness doesn't seem to hinge on the experience being a *pain* rather than some other kind of experience with a different phenomenal character.[11] I can just as easily describe a *strange*

bring out a problem that's *internal* to conceptual dualism, and I take myself to be using these terms in a way that aligns with their usage in the relevant literature.

[11] 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

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).

These each count as particular instances of what I'll call **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 at all and generating a corresponding strange experience by simply 'subtracting' the sorts of functional traits that would make it a-consciousness. @framing

Try it yourself: take any aspect of you current experience and think about what it would be like to continue enjoying that aspect of experience while subtracting any of the functional connections that put it into contact with the rest of your mind.

Since any p-conscious state can be made strange, we should be able to 'strangify' 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 articulate 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.)

It's worth examining more closely the strangest of these strange cases, since it draws the sharpest strange/non-strange contrast. What would a strange subject be like?

Strange subjects are the opposite of (philosophical) zombies.[12] And just as you can try to conceive of your zombie counterpart, you can also try to conceive of your strange subject counterpart (or, strange twin). Whereas my zombie twin is

---

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.

[12] 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).

a physical and functional duplicate that lacks p-consciousness, my strange twin will be a phenomenal duplicate that lacks my a-consciousness.

Remember, your strange twin will, by definition, duplicate *all* aspects of your phenomenology. So if you were to peer into the phenomenal life of your strange twin, things would appear just as they appear k to you now. That is, your total p-conscious experience should be 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 we run into any difficulty in trying to conceive of our strange twins, it'll have to come from an attempt to reconcile this phenomenology with a ban on a-consciousness. Again, strange subjects will share their phenomenology with us, but we're stipulating that none of their p-conscious states are a-conscious/causally connected to any of your twin's other mental state. So to conceive of your strange twin is just to conceive of subtracting a-consciousness from your own case–while leaving all of your phenomenology intact.

Can that really be done? At this point, we can identify few different ways to worry 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, perhaps various experiences have to be functionally integrated in order to form a single cohesive subject. In this case, experiences might be said to bear unity relations to one another. And it may also be said that the subject bears an ownership 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 typically think about such minds as capable of mental action. Minds are *active*, they *do* things, they make decisions, they form judgments, they initiate action, 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 decide to 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 we directly experience our own cognitive processes and mental activities. Could this sort of experience ever be illusory? And if so, are we only ever *indirectly experiencing* the effects of 'mental

action' without *directly 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 extended through time. 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/a-consciousness)? 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 each of these worries to the same degree. (I feel the pressure of the subjective unity worry much more than the other two.) But any of these worries (or others that I haven't considered) can provoke the intuition that conceiving of a phenomenal mind already involves the conception of some kind of underlying functional structure in which various experiences are embedded and through which they bear relations to one another. What I think this brings out, if only in a rough way, is 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 structural/functional/causal lens in thinking about the collection of experiences which comprise 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, which have to be understood as a certain kind of functionally organized system.

The most essential point to appreciate about the structured subject intuition–and this is a feature of any view that questions the conceivability of strange experience cases–is the way it threatens the supposed conceptual independence of phenomenal consciousness from anything functional. To take the structured subject intuition 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 phenomenological facts to functional facts. Or, to put it another way, it means that we can learn something about the functional

structure of a subject just by learning about her phenomenology.

(Consider how weird this looks from the conceptual dualist's perspective. It'd be similar in kind to the claim that, actually, black objects can't have any shape whatsoever, but only certain shapes are conceptually compatible with the color black.)

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 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 profile (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, instantiate the functional profile of some physical system. And seeing that this is a conceptual possibility can help make clear 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 which for whom its an experience?

- *The Local Mental Action Worry:* How could a strange experience play any role in mental activity?

- *The Local Temporal Relations Worry:* How could a strange pain be part of any temporally extended causal chain or endure through time?

But any underlying uneasiness is, I think, better 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 goes 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 *entirety* of one's experience. And while we may find clever ways to convince ourselves that 'hidden' phenomenology is okay in particular instances, it's much harder to make it seem okay that *all* of our phenomenology could be 'hidden'.

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 of these cases make you pause, it's worth figuring out whether or not that means you are implicitly committed to some a priori functional constraints on experience. With the structured subject intuition, I tried to provide an initial sketch of what might motivate an alternative to conceptual dualism. That alternative ought to refined and developed into a more careful philosophical position, and I pursue that project elsewhere. But the particulars of that positive account need not be settled in order to make a negative case against conceptual dualism.

4 Can Strange Experiences be Made Palatable?

Strange experiences present conceptual dualists with a challenge. If strange experiences are inconceivable, as I suggest, then conceptual dualism seems to lose its justificatory support. Without a positive conceivability verdict for cases of p-consciousness without a-consciousness, it's unclear whether the corresponding a priori entailment fails to obtain. Consequently, it's unclear whether there's really a clean distinction between thinking of consciousness in a how-it-feels sense and thinking of consciousness in a how-it-works sense. To block this line of reasoning, conceptual dualists need to explain what's mistaken in skepticism about the conceivability of strange experiences.

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 this 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 sort of approach pretty clearly can't be used in the present context, in support of conceptual dualism. It proves too much: it undermines the theoretical framework that was used to articulate conceptual dualism in the first place. Conceptual dualism *needs* the conceivability tests described earlier (and the related more-or-less-standard understanding of the a priori) in order justify the claims that make up the view.

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 the 'correct' 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, seems to pursue this kind of strategy in some places:

> "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."

Any such attempt to shift our notion of a-consciousness will have the same basic problem: it merely relocates, rather than removes, the conceptual dualist's 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 constraints. 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 misunderstand the structure of the present issue.

This maneuver is akin to pointing to some shape property–say, having some right angles–that blackness *does* a priori entail, and using that to get around worries over whether blackness a priori entail squareness. Conceptual dualism is interesting and significant insofar as it rules out a priori entailments between p-consciousness and *any* functional properties.

(An additional note: In my view, the accessed/accessible distinction is not very useful and somewhat confusing because 'access' is not a particular kind of use. If we understand a-consciousness as conferring a special status that allows a mental state to be used for a wide range of purposes, then it's unclear what the contrast between *potentially* a-conscious states and *actually* a-conscious states comes to.)

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 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.

We should be able to disarm this sort of maneuver with some of the insights gleaned from the above discussion of the accessed/accessible distinction. 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. While I don't find the accessed/accessible distinction all that helpful, it's clear that the class of states we're trying to target with 'a-consciousness' are the very same states others have tried to target with the 'accessible' label. So this sort of objection will fail for the same reasons the earlier objections failed: it doesn't address the claim that p-consciousness is tied to the functional notion of accessibility.

(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.)

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 the any act of *conceiving*. 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 'experience' concept can be applied without actively imagining or conceiving of any *particular* kind of experience. This follows from a general point about our 'experience' concept briefly gestured at earlier: 'experience' is a determinable, not a determinate. 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, to my lights, just as intuitively 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

[Draft Note: I'm not sure I should include this objection.]

A conceptual dualist might want to grant that there's something difficult about conceiving of strange experiences, but locate 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 in 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. (And so this objection may just collapse into the 'bare denial' strategy, outlined below.) 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 seem 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 knowing that you're having experiences that are unknowable.

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 fist 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.

4.6 Constraints on Phenomenal Subjects, Not Phenomenality

[Draft Note: I'm not sure I should include this objection.]

Objection: "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 way of arranging phenomenal experiences implicates functional facts. And so it's not at all clear that we can't conceive of phenomenal properties and functional properties independently of one another."

To this, I have two replies. First, whether some set of phenomenal experiences are unified together in, or belong to some phenomenal subject should, I think, count as a phenomenal fact. 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 phenomenal subject" a priori entails "S is a-conscious", I'll have accomplished my primary goal. 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 some subject. This is a familiar point, and it has been made a number of times, in different ways [CITE]. But to put it simply, many find it plausible that: if x is p-conscious, there must exist some S which is p-conscious of x. I won't mount a defense of this claim in this paper, but if it is right, then a functional constraint on phenomenal subjecthood would, consequently, constrain the entire phenomenal domain.

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. This is 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."

While I do want to acknowledge this sort of flat-footed objection as a possible view, I don't myself take it very seriously. And I'm hoping that you'll better understand my position if I run through an argument that's structurally similar to mine: Hilary Putnam's super spartans/X-worlders argument against analytical behaviorism (1963).

Putnam's argument works by carefully demonstrating the full range of consequences that follow from accepting analytical behaviorism. The hope is that, when this is all laid out clearly, analytical behaviorism no longer looks so attractive. Because we have prima facie commitments to the plausibility of super spartans/x-worlders (beings who experience pain, but behave as if they're not in pain), and because analytical behaviorism directly entails the implausibility of these prima facie plausible scenarios, the view, itself, became implausible.

Note that there's no meaningful sense in which analytical behaviorists were *forced* into admitting defeat in the face of such a challenge. Analytical behaviorism can be made to accommodate the necessary consequences, but such attempts, to an outside observer, are bound to appear ad hoc, motivated out of desperation to save the view.

So the strategy behind Putnam's argument, and mine, is: start with some judgments that are treated as certain, and build a theory on the basis of these certain commitments. Those who disagree with these certain commitments are wrong 'from the start', and their mistake can't be explained by appealing to some independent source of justification. There *is* no such independent source of support, nor *can* there be. So long as we've done out best to rule out conceptual confusion, any leftover disagreement can only be the result of a fundamental, ground-level misunderstanding.

But there's one important disanalogy between Putnam's argument and mine: unlike analytical behaviorism, conceptual dualism is frequently taken as an uncontroversial background commitment in philosophical discussions of consciousness. So while a stalemate might not have represented progress for the opponents of analytical behaviorism, I hope a stalemate, here, will represent a kind of progress–if only for contingent, historical reasons.

By reexamining conceptual dualism and dislodging it from its privileged position—by foregrounding an issue that has 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 inductive 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, it'd be a mistake to get hung up on this leftover task, since the basic idea that there are functional constraints built into the notion of experience already has some interesting consequences.

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 though/talk can be *reduced* to functional though/talk, the view in question only claims that the determinable concept of 'experience'/'p-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. (I, myself, would reject the possibility of such an exhaustive analysis, but this possibility is left open by the view.) 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 priori, a-conscious, it does not tell us anything about the functional profile that distinguishes pain experience from other kinds of experience.

Second, 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. (I, myself, would reject any functional-to-phenomenal entailments, but the view I'm laying our requires no such commitment.)

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 'experience' concept properly applies).

That said, a revision to our understanding of the 'experience' concept will, necessarily, trigger revisions to our understanding of 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 a science/metaphysics of mind. For example, we might be able 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 also might 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.

Whether or not such a view can really be developed into full-fledged philosophical position remains to be seen. 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 view.

6 References

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.

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.

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

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

Putnam, Hilary (1963). Brains and behavior. In Ronald J. Butler (ed.), Analytical Philosophy: Second Series. Blackwell.