Dennettian Panpsychism: Multiple Drafts, All of Them Conscious

Luke Roelofs[1] ⬤

Abstract

I explore some surprising convergences between apparently opposite theories of consciousness—panpsychism (specifically constitutive panpsychism) and eliminativism (specifically Dennett's multiple drafts model). I outline what a 'Dennettian panpsychism' might look like, and consider some of the challenging but fertile questions it raises about determinacy, holism, and subjecthood.
What unites constitutive panpsychism and the multiple drafts model is that both present the unitary consciousness we can report as resting atop a multiplicity of independent processes; both reject as misguided the search for a definite threshold between processing that is truly conscious and that which is merely preconscious. What divides them is that Dennett regards it as unreasonable to posit inaccessible consciousness, but reasonable to doubt or deny the existence of consciousness, while panpsychists think the opposite.

Keywords Consciousness · Panpsychism · Eliminativism · Multiple drafts · Dennett · Cognitive science

It might seem like eliminative physicalists, who say nothing is conscious, are the polar opposites of panpsychists, who say everything is conscious. But in some ways, the two positions are closer than they appear. Both reject the attempt to draw a boundary between systems or events that have this special thing called 'consciousness' and those that don't—a rejection particularly marked in the work of Daniel Dennett, whose 'multiple drafts' model is generally taken to effectively eliminate consciousness by deconstructing the idea of a transition from 'pre-conscious' to fully 'conscious' processing. Dennett accomplishes this deconstruction in part by focusing on the multiplicity of processing streams, and the clash between this multiplicity and our introspective impression of a clear and unified conscious field. This focus on real multiplicity and apparent unity may make the multiple drafts model

✉ Luke Roelofs
  luke.mf.roelofs@gmail.com

1   New York University, New York City, NY, USA

⦚ Springer

surprisingly attractive to constitutive panpsychists, who regard human consciousness as built up from fundamental consciousness through a combination of many simpler conscious parts. My aim in this paper is to explore the prospects for a view that combines the multiple drafts model and constitutive panpsychism: Sects. 1 and 2 outline what these two ideas are individually, Sect. 3 explains what their combination looks like, Sects. 4 and 5 consider some of the difficult but fertile questions faced by that combination, and Sect. 6 closes with some reflections on philosophical methodology.

1 The Multiple Drafts Model

Here is how Dennett introduces his model
According to the Multiple Drafts model, all varieties of […] mental activity […] are accomplished in the brain by parallel, multitrack processes of interpretation and elaboration of sensory inputs…once a particular 'observation' of some feature has been made, by a specialized, localized portion of the brain, the information content thus fixed does not have to be sent somewhere else, [for] re-presentation of the already discriminated feature for the benefit of the audience in the Cartesian Theater–for there is no Cartesian Theater.
These spatially and temporally distributed content-fixations in the brain are precisely locatable in both space and time, but their onsets do not mark the onset of consciousness of their content. It is always an open question whether any particular content thus discriminated will eventually appear as an element in conscious experience, and it is a confusion, as we shall see, to ask when it becomes conscious. (Dennett, 1991, pp.111-113)

A key element of this proposal is the denial of a precise, objective moment in time or location in space where non-conscious content becomes conscious: there is no 'magic threshold' in the brain that delimits the conscious states from the non-conscious ones.[1] Rather, Dennett suggests, there are many ongoing processes of content-fixation, operating in parallel and on slightly different timelines, taking in input and passing on their output to other systems. There is no 'Cartesian theatre' where they have to all send their decisions to be 'displayed in consciousness', but neither are they themselves sufficient to provide a determinate, consistent, conscious experience, because their contents often contradict each other, or are in the process of revision. Rather than a single story, there are 'multiple drafts' being written and

---

[1] For my interpretation here, I am indebted to Akins (1996), who pulls out a number of key theses that constitute the multiple drafts model: the three most important for the discussion in this section are:

'There are no 'fixed facts' about the events and contents of conscious experience other than facts about probes (and bindings). […].

On a small time scale, there is no distinction, in principle, between a revision of perceptual input and a revision of memory. […].

There is no determinate time of occurrence for conscious events. (Akins 1996, pp.2–3).

revised in tandem. On this view, our impression of a single, consistent, determinate experience at any given time is a product of a process (which Dennett leaves a little under-theorized) of 'probing' these multiple streams: we, or someone else, 'poses a question' and in the course of generating an answer, the brain settles on a definite verdict from each subsystem and binds them together into a definite, consistent, answer.

Dennett (1991, pp.114–126) highlights this model's capacity to account for a certain sort of illusion where two stimuli appearing at different locations in rapid succession give rise to an experience of a single object moving from one location to another. He discusses particularly the 'color phi' phenomenon (Kolers & von Grünau, 1976; Wertheimer, 1912), where flashing someone a red dot (say) on their left followed quickly by a green dot on their right gives them the impression of a single dot moving while changing colour, and the 'cutaneous rabbit' (Geldard, 1982; Geldard & Sherrick, 1972), where taps at the wrist and the elbow give rise to an experience as of multiple taps 'hopping' along the forearm, even though the forearm was not touched.

These cases involve perceiving a stimulus which was never there, but that is not what makes them interesting. What makes them interesting is that the perception of the illusion seems to happen *before* the perception of the later real stimuli, even though the latter is necessary for the illusion to take place at all. It seems that the brain must detect the second stimulus, form a 'best guess' model of what caused it, and 'fill in' an experience of the illusory intermediary, but must do all of this *before* there is any conscious experience of the second tap. Of course, it is not strange for there to be some time between us receiving a stimulus and undergoing an experience. But the length of the gap in this case is implausibly large, Dennett thinks, because it exceeds the gap involved in constructing consciousness experiences in other, simpler. He notes that we might then be tempted by a different picture: the experience of the final stimulus happens quickly, but as soon as some brain mechanism computes that the most likely cause of two such rapid stimuli is a single moving object, there is a swift revision of memory, so that the subject falsely remembers having experienced the intermediate stages of the movement.

Dennett terms these models the 'Stalinesque' and 'Orwellian,' since the one constructs a 'show trial' (a false perception) while the other revises 'historical records' (a false memory). But instead of arguing for one over the other, he presents them as both implausible, so that the choice between them is a sort of reductio ad absurdum of the assumptions that lead us to it. The Stalinesque model would have the brain 'waiting' to experience the second stimulus before experiencing anything, which seems senseless for an organ that evolved for speed and efficiency. On the other hand, the Orwellian model posits an equally senseless capacity for making rapid, undetected, revisions to memory records. But if we reject the idea of a 'Cartesian theatre', somewhere in the brain where conscious experience is 'played' for consciousness, we don't have to choose either of the two. Rather, Dennett suggests, it is simply that different subsystems are processing different aspects of the stimulus in parallel, not waiting for one another but updating each other with their results as soon as those results are produced, and revising their results in light of those from other subsystems. These are the 'multiple drafts': an ongoing flow of shifting,

incomplete, stories of what happened. In the cutaneous rabbit, for instance, there is first a quick determination of certain features of the second stimulus, followed by a determination that the most likely cause was a single moving object, leading to a representation of intermediate stimuli that arises after the representation of the second actual stimulus, but is represented as occurring before it. There exist simultaneously two 'drafts' of what was perceived: one featuring fewer taps, the other featuring more.

These illusions show a case where different brain systems determine different aspects of the same content at different times. But Dennett suggests that this is likely to be true much more generally, and not unique to these illusory cases. After all, from an evolutionary perspective, a brain which does not force all subsystems to wait for the slowest one will react faster, which might make the difference between life and death. If a particular object's colour can be determined by dedicated colour-processing systems before its shape can be determined by shape-processing systems, it makes sense to initiate colour-based responses—potentially including report, reflection, memory—as soon as possible, rather than to wait around for shape to be determined as well. But accepting the likelihood of the brain working this way raises difficult questions about consciousness. Is the subject conscious of the colour immediately, or only later when they can be conscious of shape as well? Why do they only ever seem to experience shape and colour together?

The most radical feature of Dennett's proposal is his denial of an objective fact about exactly what the subject is conscious of at any particular moment, independently of particular probes. As he puts it:

> ...the fundamental implication of the Multiple Drafts model [is] if one wants to settle on some moment of processing in the brain as the moment of consciousness, this has to be arbitrary. One can always 'draw a line' in the stream of processing in the brain, but there are no functional differences that could motivate declaring all prior stages and revisions to be unconscious or preconscious adjustments, and all subsequent emendations to the content (as revealed by recollection) to be post-experiential memory contamination. The distinction lapses in close quarters. (1991, p.126)

That is, among all the multiple, sometimes conflicting, content-fixations going on in different subsystems, there might be multiple equally good candidates for 'what is experienced'. Their interactions—checking for consistency and revising what is found inconsistent—are not inherently either pre-consciousness ('pre-publication edits') or post-consciousness ('post-publication revisions'), because there is no moment of publication. There are just the multiple drafts, churning along together. On this account, it only seems to us that we have a single consistent timeline of consciousness because if someone asks us 'What did you experience?', our various subsystems will settle on mutually consistent contents, a single coherent narrative, to guide the behaviour that answers the question (to be reported, remembered, or otherwise made use of). This seems to imply a sort of eliminativism about consciousness: if 'there are no fixed facts about the stream of consciousness independent of particular probes' (Dennett, 1991, p.138), then there isn't really such a thing as consciousness, at least as we normally think of it.

It might seem that consciousness can simply be identified with the response to a probe, and at times Dennett seems to lean towards this position. But this requires giving up the seeming truism that we are usually conscious throughout the day from when we wake up to when we fall asleep—since on this reading of the multiple drafts model, during any time when we are not probing our stream of consciousness, it does not exist! That is, if we are not attending—not asking ourselves 'what am I experiencing right now?'—we are experiencing nothing. This clash between the sparseness of probes and the constancy of what we think of as 'consciousness' militates against identifying consciousness with the responses to probes. Instead, it supports identifying 'consciousness' with the ongoing streams of perceptual determinations that sometimes get probed. But conversely, the indeterminacy and inconsistency of these streams militates against identifying them with 'consciousness', yielding the result that no single thing in the brain is a good fit for our everyday notion of 'consciousness', and that to that extent 'consciousness' does not exist (cf. Dennett, 1988).

2 Constitutive Panpsychism

The denial that consciousness exists is a hard pill to swallow. For those who are attracted to other aspects of the multiple drafts model, but find its eliminativism off putting, it might be useful to explore how much of the model can survive with the eliminativism eliminated. The natural way to do this is to suppose that consciousness is 'spread out' among the multiple drafts in their ongoing, messy revisions and bindings: consciousness overflows the polished drafts that we report when probed. The problem with this, from Dennett's perspective, is that it breaks the link between consciousness and reportability, since report always takes in only a tidied-up subset of what is really conscious. But if consciousness cannot be identified with what the subject reports, what is it? And where is it: when does consciousness first appear, and how can any moment of first appearance be made empirically meaningful, if it is beyond the reach of subjective report? Is there consciousness in the optic nerve? In individual cells? Does consciousness go all the way down?

To Dennett, the empirical intractability of these questions constitutes a reductio ab absurdum of spreading consciousness more widely. But some philosophers think we already have good, independent, reasons to think that consciousness goes all the way down: reflection on physics, phenomenology, and what they can or can't ever say to one another pushes us towards 'panpsychism', the belief that consciousness must be an inherent, fundamental feature of matter generally.

Recent work on panpsychism has not tended to engage with the details of cognitive science work on consciousness; rather, it has focused on the possibility in principle of explaining consciousness reductively, and the implications of different non-reductive models (see, e.g. Nagel 2004; Chalmers, 1995, 2015; Strawson, 2006; Goff, 2013, 2017; Mørch, 2014). To summarize very roughly, contemporary panpsychists think that the various arguments brought under the heading of the 'hard problem of consciousness' provide a decisive objection to physicalism: no amount of purely physical information can explain why any physical process is conscious.

Consciousness, they conclude, cannot be accounted for in terms of wholly non-conscious ingredients: a world built from mindless parts would itself be mindless. It seems to follow that the presence of consciousness in our world reveals something about its fundamental constitution: there must be some sort of basic, fundamental consciousness somewhere, since without that, consciousness could not arise. One way this might be is if some sort of strong emergentism or dualism were true: *most* of the world is exhaustively described in physical terms, but consciousness is an extra fundamental property that appears specifically in human and animal brains. But panpsychists reject this sort of antiphysicalism that would separate the conscious mind from the rest of nature: instead, they argue that since we each know that at least one physical system—our own brain—somehow involves consciousness, and since all physical systems are alike in their fundamental constituents, the most likely option is that all physical systems are in some way like our brains, involving consciousness with structure and complexity proportional to their observable physical structure and complexity.

Obviously, this argument is far from uncontroversial, and my aim here is not to attack or defend it. I wish instead to draw out a neglected affinity between panpsychism and its seeming opposite, eliminativism. The affinity comes from the fact that many panpsychists take an approach known as 'constitutive' panpsychism, on which human consciousness arises from 'combination' of the simpler conscious states associated with the simpler parts of the brain, ultimately all the way down to fundamental particles. This contrasts with what is often called 'emergentist' panpsychism, on which human consciousness is a metaphysical 'jump' from its conscious components, requiring some sort of special metaphysical bridge from micro to macro. Of the two, constitutive panpsychism is often thought more attractive *if* it can be made to work, but to face more serious versions of the 'combination problem', the difficulty of making sense of how consciousness combines (cf. Seager, 1995, p.280, Chalmers, 2017, Montero, 2017, Miller, 2017, 2018, Roelofs 2016, 2019, 2020, Mendelovici, 2020). That is, constitutive panpsychism requires the possibility of multiple conscious states coming together somehow into a single conscious state—requires, in short, that our conscious states arise from multiple more basic states.

This makes it useful to see how far constitutive panpsychism can take over the resources of structurally similar non-panpsychist theories, such as Dennett's. If consideration of phenomena like the cutaneous rabbit has led Dennett towards seeing everyday consciousness as a messy collation of multiple drafts, this may indirectly help constitutive panpsychists who are pushed in similar directions by very different arguments.

3 Multiple Conscious Drafts

What would a panpsychist version of the multiple drafts model look like? It would agree with Dennett that multiple drafts are formed and revised in parallel, and that there is no magic threshold in this process where consciousness appears. But rather than saying that really, none of this is conscious, or that only the reportable

outcomes of probes are conscious, it would say that everything—all of the multiple drafts—are conscious. Each local content-determination by a different neural subsystem is a phenomenally conscious event, and the constant mutual comparison, correction, and integration among subsystems involve phenomenal relations among these phenomenal events. Strictly, it's not even just that each subsystem involves consciousness; neural subsystems are themselves composed of smaller parts, down to individual neurones, then to individual molecules, ions, and so on down to the most fundamental level, if there is one. How all of these layers of micro- and macro-consciousness relate is a major question for panpsychists; the most natural line for someone trying to develop a Dennettian version would be that something like the relation between whole-brain-consciousness and neural-subsystem-consciousness also obtains between the latter and sub-subsystem-consciousness, iterating at every step. That would mean that what we normally think of as 'our consciousness' is just the tip of the iceberg, and moreover that the local content determinations discussed by Dennett are the parts of the iceberg just below the waterline. I'll come back to how to think about the further-down parts of the iceberg in Sect. 5, but for now, let's focus where Dennett does, on the relation between the top level and what is immediately below it.

Dennettian panpsychism says that the pattern of phenomenal relations among phenomenal elements in subsystems plays a crucial role in cognitive accessibility: cognitive access ('probing') by the rest of the system both facilitates and relies upon the different fragmentary consciousnesses organizing themselves into a coherent structure. That is, the familiar sort of coherent, determinate, experiences that we attend to and report are what happens when these multiple local consciousnesses get themselves into good order and 'agree on a story'. But even when they are not in good order, they are still conscious, and as a result, this model would make our conscious state 'fuzzy around the edges': some of its elements, especially those at the periphery of attention, are not in stable, consistent, relations to one another.

In the next section, I'll examine these phenomenal elements and phenomenal relations a little more closely. But first, it is worth getting clear on what is distinctively Dennettian about Dennettian panpsychism. Many forms of panpsychism accept the idea of phenomenal elements, bound together by phenomenal relations (though some do not), so that is not very distinctive. We can note that Dennettian panpsychism is necessarily a constitutive form of panpsychism, rather than a strongly emergentist form. Emergentist panpsychism involves precisely the kind of discontinuity Dennett seeks to avoid between 'my consciousness' and the various elements that give rise to it. But again, rejecting strong emergence is not particularly distinctive. I think the distinctively Dennettian flavour is best seen by considering a contrast between different ways of developing constitutive panpsychism: the multiple drafts direction considered here, and a contrasting direction more akin to what Dennett decries as 'Cartesian materialism'. The latter direction would involve saying that there is some sort of minimum necessary degree of phenomenal binding or integration, such that only when the component experiences meet that threshold does the whole human subject themselves start to experience them. On this view, there is still a 'magic threshold' between conscious-for-me and not-conscious-for-me, even if the events that are not conscious-for-me, because of being insufficiently integrated, are

nevertheless conscious in and of themselves. There is still, thus, a sharp boundary to 'my experience', and a corresponding distinction between a 'prepublication' change to perceptual content, and a 'post-publication' revision of memory[2].

This more 'Cartesian' sort of panpsychism would have something common with Zeki's (2007) suggestion that multiple brain areas independently generate 'micro-consciousnesses' for different perceived qualities, which are only secondarily unified into 'macroconsciousness'. Importantly, Zeki's view is that upon being unified together, the microconsciousnesses cease to exist in their own right (p.584), while constitutive panpsychism can allow them to continue to exist as components of the macroconsciousness. That is, constitutive panpsychism usually allows for 'experience-sharing' (cf. Roelofs 2020; Goff and Roelofs Forthcoming): a particular phenomenal component, like the registering-as-red that might occur in one cortical area, can belong simultaneously to multiple overlapping phenomenal wholes, both the overarching thing I call 'my consciousness' and the various 'drafts' that are woven together to form it. This means that constitutive panpsychists, unlike Zeki, don't have to posit a single decisive transition when we cross the threshold from 'micro' to 'macro'. Indeed, they need not posit any such decisive transition, however understood. 'Dennettian panpsychism' is what you get from consistently embracing this rejection of thresholds: it is constitutive panpsychism which resolutely refuses to draw a sharp divide between what is and what is not part of 'my' consciousness. It contrasts with forms of panpsychism (constitutive or emergentist) which do draw such a divide, forms which we might consider calling 'Cartesian panpsychism' by analogy to Dennett's term 'Cartesian materialism'.

4 Difficult Questions

So the distinctive commitment of Dennettian panpsychism, as I am understanding it, is the claim that even though the binding relations among different subsystems are messy, unstable, incomplete, and sometimes contradictory, that does not stop us being conscious of them. It just means that our stream of consciousness is messy, unstable, incomplete, and sometimes contradictory. In this way, the panpsychist can follow Dennett in avoiding the awkward choice between Stalinesque and Orwellian models of illusions like the cutaneous rabbit. The multiple drafts theory says that there is no objective fact about whether the illusory intermediate touches are inserted into the story 'before consciousness' (Stalinesque) or 'after consciousness' (Orwellian), because there is no objective fact about what is and isn't conscious. Dennettian panpsychism says that it is all conscious, equally, but forms a messy composite stream of consciousness, with fuzzy edges and indeterminate

---

[2] On 'fusionist' panpsychism, defended by Seager (2010, 2017) and Mørch (2014), the phenomenal field of a human being is an indivisible whole, which arises when many simpler phenomenal elements come together in the right configuration but is not in any literal sense composed of those elements: once the whole exist, the parts are gone. A similar mutual exclusion between conscious parts and conscious whole is built into the Integrated Information theory (Oizumi et al., 2014; Tononi 2012), which is not panpsychist but close to it.

content. But this brief description raises some difficult questions: what exactly are these phenomenal parts, and the phenomenal relations between them? What does it mean for them to form a stream with fuzzy edges and indeterminate content? Let me say a little about the difficulties raised by these questions.

4.1 Phenomenal Elements

First off, what exactly is the Dennettian panpsychist's picture of these local content determinations—these component consciousnesses? They must have something in common with aspects of familiar complex conscious states: when brain area MT detects motion with a certain speed and direction, this is phenomenologically something like when we see a coloured, shaped, object at a location moving with that speed and direction. Yet it is also different, because it does not involve any experience of colour, shape, or location. One worry is that this sort of experience is simply impossible: how could there be motion without change of location, or without any visible thing that moves? One reply is that the waterfall illusion seems to involve precisely the experience of motion without change of location (see Crane, 1988), but a more fundamental reply is that while visual motion all by itself may be unimaginable to us, that is not surprising—we are complex beings whose visual system compulsively binds together visual motion with other features, and so of course when we try to imagine motion we struggle not to also imagine other features. But just as the unimaginability of bat experience to a human being (or vice versa) does not suggest its impossibility, so this does not suggest that visual motion experienced by itself is impossible.

Someone might raise a concern here about the 'interpenetration' of different experienced features. It might seem like the experience of visual motion, say, is so intimately connected to experienced colour, shape, and so on that they change its intrinsic character, so that an isolated experience could not have the very same quality—motion-in-isolation and motion-of-a-coloured-shape are two different phenomenal characters (cf. Fink et al. Forthcoming). This is an especially pressing concern for panpsychists, since they are often committed to the idea that consciousness is intrinsic: a conscious event or state has its particular phenomenal character in itself, not in virtue of anything outside itself. This might seem to be at odds with the phenomenological observation of interpenetration.

One option, in response, would be to question whether there really is such a deep interpenetration: of course experiencing motion and colour is different from experiencing motion without colour, but does it follow that the experienced motion *itself* is different in the two cases? Is there really no common element? But a more robust, and more interesting, response is that the Dennettian panpsychist can accommodate interpenetration, because they can allow that the phenomenal relations among these components will causally alter their intrinsic phenomenal character. Activity in area MT should not be thought of as an experience of motion-in-isolation, because that activity is not isolated: it is in constant causal communication with activity elsewhere, and this will change its intrinsic character in subtle ways. This might involve giving it the 'interpenetrated' character of presenting 'motion of a coloured

shape'—albeit without presenting any specific colour or shape. What is provocative about this proposal is that the experience of MT still does not involve any presentation of colour or shape: it is just an experience of motion, but an experience with the *intrinsic* character of 'motion-as-of-some-coloured-shape', with colour and shape at most implicitly co-present through the experienced motion 'pointing outside itself' towards other phenomenal events.

This response is, as I say, provocative, but it is not incoherent—a claim I defend more fully elsewhere (Roelofs 2016, pp. 3210-3218, 2019, pp.57 ff). But since it places such emphasis on the impact of phenomenal relations, which transform the intrinsic character of individual experiential elements so that they reflect other elements, we should examine what can be said about these relations.

4.2 Phenomenal Relations

So what exactly are the phenomenal relations posited by Dennettian panpsychism? Dennett and Akins refer to the relations among different content determinations in various ways, but the most prominent term is 'bindings'—the colour identified in one brain area, the shape identified in another, the motion identified in a third, are all 'bound together' somehow (cf. Duncan & Humphreys, 1989, 1992; Treisman & Gelade, 1980; Treisman & Sato, 1990). The Dennettian panpsychist can say basically the same—the crucial relation is one which builds, for example, 'experience of a red triangle' out of 'experience of red' and 'experience of a triangle'. What the panpsychist can add is that this binding is a specific form of the relation often called 'phenomenal unity', 'co-consciousness', or 'experienced-together-ness' (cf. Bayne, 2010; Bayne & Chalmers, 2003; Dainton, 2000; Roelofs 2016, Schechter, 2013; Shoemaker, 2003). Phenomenal unity in general is the relation that obtains between two experiences when there is something it is like to experience both together; what we can call 'phenomenal binding' is the relation that obtains between two contentful experiences when experiencing them both together involves experiencing their contents as standing in some specific relation (for more discussion of phenomenal binding, see Roelofs 2019, pp.170-182, cf. Woodward, 2015, 2021).

One notable consequence of Dennettian panpsychism is a considerable degree of 'indeterminacy' in human experience. That's because on this view, the process of making particular experiences cognitively accessible is a process of aligning them with other experiences so that a single, determinate, consistent picture can 'coalesce', and this process will happen by degrees over a period of hundreds of milliseconds, constantly updating as new information comes in. At many points, there will be no objective fact about whether a given element is included in a given 'draft', or whether a given 'draft' is integrated enough, internally or with other drafts, to be called 'what I am experiencing' in the everyday sense. And if two subsystems have conflicting contents (e.g. 'taps all along the forearm' and 'taps only at the wrist and elbow') then, until one or the other is revised, the total stream will include both, thereby representing indeterminately whether there are or are not taps along the forearm.

Let me say a bit more about this sort of indeterminacy. What I mean is that two or more inconsistent contents are equally good candidates for being assigned as 'how things seem' to the subject. This might be because of multiple conflicting stories about a single feature (e.g. conflicting numbers of taps) or because although each individual subsystem has settled its story (colour system says 'red and green', shape system says 'square and triangle'), there is not yet a relation between the systems that binds them (specifies which shape has which colour) and so multiple conflicting bindings are equally live options (both 'red square and green triangle' and 'green square and red triangle'). It is a little like the situation with borderline cases of vague predicates, where two incompatible contents (e.g. 'this person is bald' and 'this person is not bald') are equally live descriptions, though of course the reasons for the indeterminacy are very different, stemming from facts about our neural architecture rather than facts about our concepts and classifications.

Talk of indeterminacy makes sense when we are discussing content; at a functional level, the description would be more like 'an unstable state where two mutually-inhibitory patterns, with divergent effects on downstream processing, are simultaneously active'. The further claim of indeterminate content will only apply to the extent that we think of those neural patterns as having representational content.[3] But alongside the representational and functional descriptions, there is also a question of phenomenology: what is it like to phenomenally represent an indeterminate stimulus?

It is hard to give a fully general answer, but there do seem to be a number of instances available to reflect on. Plausibly, much of peripheral vision (and perhaps unattended or diffusely attended perception generally, and perhaps imaginings and dreams) has a significant degree of indeterminacy to it. In these areas of our experience, there might be determinations of, for instance, particular experienced features but without any determination of how they relate—there might be colours and shapes without any fact about which colours belong to which shapes. This can at first sound bizarre, but I think it actually accords fairly well with studies showing that peripheral vision, for instance, has remarkably poor discrimination ability (see, e.g. Pelli & Tillman, 2008, Levi, 2008, Freeman & Simoncelli, 2011), and with the persistence of debates about cases where we seem to determinately experience something as 'speckled' or 'striped' but not as having any determinate number of speckles or stripes (see, e.g. Tye, 2009, 2010, Fink Ms., cf. Morrison, 2016; for more details see Roelofs 2019, pp.168-170.)

Moreover, the multiple drafts suggest that often, when a new stimulus arises suddenly, the most accurate description of what we experience at first may be 'something', i.e. a highly indeterminate content, when it has been registered that something is happening but not yet determined what. This seems to me plausible: new

---

[3] Dennett talks about the multiple drafts almost always in representational terms: he is interested in what seems to the subject to be happening, how the world is presented to them. Most panpsychists are not pan-representationalists: that is, they generally do not take all forms of consciousness to have representational content. This means that the account of determinacy and indeterminacy given here can't go 'all the way down' to the level of molecules and atoms. But it is perfectly compatible with this to think of consciousness as mostly or always representational once we reach the level of neural subsystems.

experiences do seem to appear as coalescing, a rapid focusing and reduction of initial blur, rather than as a definite image appearing out of nothing. This idea is not exclusive to any sort of multiple drafts model (Husserl's account of protentions and retentions suggests something like it, cf. Husserl, 1991), but if it is right, then it can be added to cases like peripheral vision and indeterminate-speckle-number, to support the idea that pervasive fuzziness around the edges of our experience is intelligible and plausible.

5 Who is Experiencing?

One obvious question looms, regarding the component experiences associated with neural subsystems: who is having them? And if there are, as the panpsychist might think, further component experiences further down, at a yet smaller scale, which have to be woven together to form the experiences associated with a subsystem, who is having those? Here, there are some difficult choices for the panpsychist.

One option is to say that there is just one subject, the familiar one I call 'me'. Each of us is having all the experiences based in our neural subsystems, as well as the complex experience they all form—indeed, is having those component experiences precisely insofar as they have the complex experience they all form. And if those component experiences themselves arise from a process of binding together even simpler experiences (as constitutive panpsychists think), then perhaps we are having those even simpler experiences too—perhaps even all the way down to the microexperiences of our subatomic parts. This would be a lot of experiences! But in another sense, it might be no more experiences than we already thought we had—if all of these many simple experiences really are just components of the complex experiences we are familiar with.

Considered as individual experiences, the components postulated by Dennettian panpsychism are largely not cognitively accessible by the person themselves: cognitive access goes hand-in-hand with binding, and so whatever we access tends to be many components bound together. This applies even more strongly to whatever sub-components the components have: if area V4 determines a particular colour by binding together sub-colour ingredients in a particular way, it seems that we have no direct access to these sub-colour ingredients, except insofar as they are bound together. This indissociability is a major limitation in our access to these experiences, even if they are experiences and belong to us. A contrasting limitation is that insofar as component experiences are not bound with others, they are cognitively inaccessible because they cannot be attended.

Sometimes, it is suggested that, as Kant puts it, unless I am able to self-ascribe an experience, that experience cannot be properly called mine—it 'would be nothing for me' (1999, p.246). In line with this thought, one might take the above reflections on my very limited cognitive access to the experiences associated with my neural subsystems as providing good reasons to deny that they are experiences *of mine*. Maybe I myself really only experience what is accessible, namely the consistent, determinate, parts of the whole experiential complex. Maybe the other experiences do not belong to me, but to my parts, understood as each a little sub-subjects: a 'pandemonium' architecture

with conscious demons. After all, lots of people consider it at least a live option that insects are conscious with only a million or so neurones, and our multi-billion-neuron brains have plenty of subsystems that size or larger. A third option (and the one I am inclined to favour) is that these experiences are shared: they belong not only to conscious subsystems, but also to me (the subsystems are, after all, just parts of me). Or maybe, fourthly, these experiences do not belong to anyone—maybe experiences don't always have to belong to a subject? That is, we can canvas at least four initial options for who is experiencing the many strands and fragments of consciousness posited by Dennettian panpsychism:

1. I alone experience them all.
2. I do not experience them, but parts of me do.
3. Both I and my various parts experience them.
4. No subject experiences them.

I won't try to judge between these options, in part because doing so requires grappling with some more basic conceptual questions, about the meaning of this notion 'subject', its relationship to terms like 'self', 'person', 'mind', etc., and what criteria are relevant for individuating one. But whichever way we might go on these venerable questions of metaphysics, the multiple drafts model should direct our attention towards the following awkward fact: cognitive access is not all or nothing, but a spectrum of different modes of access, to more or less information, with greater or lesser ease and reliability, under varying circumstances. Given this, any distinction we might want to draw in terms of access, between the experiences that are 'mine' and those that are not, cannot be a sharp or strict one: there will be many experiences (and, perhaps, many periods of time for each experience) that neither clearly satisfy nor clearly do not satisfy the criteria for belonging to me. This is, again, somewhat in line with Dennett's deflationary position on the 'self' (1992): fundamentally, there are just all of these many interesting events happening in different brain areas, entering into all of these many interesting relationships, and any delimitation of a single distinct 'me' who they all belong to is more like a useful organizing tool than a description of objective reality.

None of these options is especially straightforward or intuitive: they are all challenging to think about. But then, so is the multiple drafts model, and eliminativism more broadly. Part of the point of such models was that our familiar ways of thinking about the mind are likely to be unsatisfying—that whatever the truth is, it will almost certainly seem 'crazy' in one way or another (cf. Schwitzgebel, 2014). Holding onto magic thresholds is a way to keep things easy and safe, but it's a mistake. Decisively throwing out the magic thresholds, in either an eliminativist or a panpsychist way, can help us come to terms with how the brain really seems to work, at the cost of forcing us to rethink our most basic assumptions about the mind.

6 Philosophical Methodology

Beyond Dennett's specific model, its animating spirit (the rejection of the search for a determinate threshold of consciousness) can provide a useful lens for looking at many discussions in cognitive science. Consider the debate about Block's notion of 'phenomenal overflow', i.e. about whether consciousness is richer than what can be cognitively accessed (Dehaene et al., 2006, Block, 2007, 2008, 2011, Kouider and Dehaene, 2007; Kouider et al., 2012, Cohen & Dennett, 2011, Brown 2012, Haun et al., 2017), and its recent neuroscientific offshoot, the debate about whether visual consciousness depends partly on activation in the prefrontal, or just on occipital, cortex (Lau & Rosenthal, 2011, Tsuchiya et al. 2015, Lamme, 2016, Boly et al., 2017). Part of what is at issue here is simply getting a clearer sense of the sequence, capacity, and structure of various processing stages—whether, for instance, there are two stores of diminishing capacity, or a single set of probabilistic representations that support limited sampling processes (Gross & Flombaum, 2017), and how far early processing of 'gist properties' can be verbally reported (Bronfman et al., 2014). But a key part of the debate is not about this—it is about when consciousness first appears. Even if all parties agreed on the sequence of physical events from retinal stimulation to occipital activity, then to prefrontal activity, and then to report, the crucial disagreement is which stage first involves phenomenal consciousness. The multiple drafts model regards this as a bad question, and so too do panpsychists: there is always consciousness in the brain (indeed on some versions nothing but consciousness) and the question is just what forms and shapes it assumes under different conditions, which what relations to cognitive access and the identity of the (or a) subject.

Part of the debate here is whether it is legitimate at all to posit inaccessible consciousness. Dennett denies this, regarding such a posit as inherently unscientific. Panpsychists, fairly obviously, side with Block and others who think inaccessible consciousness is a legitimate posit: on their view, most of the consciousness in the universe is inaccessible, in that it exists in nonhuman or non-living systems without the cognitive structure required to speak meaningfully of 'access'. But they can agree with Dennett that its existence or non-existence can't be directly verified or falsified: experiments which seem to support positing it rely on showing dissociations between one sort of access and another (e.g. the results of Sperling 1960 and Bronfman et al., 2014, where reports of seeing many letters or colours come apart from reports of specific items), and so do not constitute *experimental* evidence that consciousness is possible without any sort of access at all.

Block defends the posit of phenomenal overflow by saying that inaccessible consciousness may be posited by inference to the best explanation: when the same brain activity, under the same stimulus conditions, occurs both with and without reportability, the simplest explanation may be that it is conscious in both cases. But what counts as the best explanation is partly a matter of what fits best with background theoretical and conceptual commitments: many of Block's critics clearly think that the default interpretation of any given brain activity is that

it is not conscious, until evidence can demonstrate its consciousness, and Cohen and Dennett explicitly doubt the very coherence of the idea ('What does it mean to have a conscious experience that you yourself do not realize you are having?' p.362). By contrast, panpsychists will take consciousness to be the default assumption, and lack of (some particular kind of) report is just evidence that it is inaccessible rather than accessible (in some particular way). It is a mistake to expect empirical data to decide between theories rooted in different conceptual and metaphysical starting points.

Cohen and Dennett draw the conclusion that 'the so-called hard problem of consciousness […] far from being a formidable obstacle to science, […] achieves its apparent hardness by being systematically outside of science' (2011, p.363). Panpsychists can agree, though again putting a different spin on the idea. They usually see the hard problem as a crucial motivation for their theories, but need not take it to be either a problem for, or obstacle to, science. That a problem is outside of science does not mean that it is pseudoscientific or antiscientific, just that different solutions to it will imply different interpretations of any scientific data or scientific theories that we develop, and so must be evaluated in part by non-empirical criteria, such as those involved in debates about the hard problem.

Panpsychism, reductive physicalism, and eliminativism are in a position somewhat like realism, idealism, and solipsism about the external world. No specific experiment in physics will reveal the right account of the external world: the choice among them is in that sense outside of science. But the choice is not therefore meaningless, and the interpretations given by one of these theories are not 'unscientific' simply because the choice between the theory they reflect and its rivals is not a scientific question. Rather than rejecting panpsychism, or other theories which posit inaccessible consciousness, as 'unscientific' for clashing with one set of metaphysical and conceptual assumptions, we should accept that scientific results will not determine a unique best interpretation as long as different broader theories of consciousness continue to be live options. New data will support different conclusions in light of different theories of consciousness, and there is little prospect of a decisive victory of one such theory over the others. For the foreseeable future, cognitive science might have to include, side-by-side, work on reductive physicalist cognitive science, on eliminativist cognitive science, and on panpsychist cognitive science, posing similar empirical questions but giving different interpretations of the answers. And while in some ways the former two are closer to each other than to the third, in other ways, eliminativism and panpsychism are surprisingly akin.

Declarations

Conflict of Interest The authors declare no competing interests.

References

Akins, K. (1996). Lost the plot? Reconstructing Dennett's multiple drafts theory of consciousness. *Mind & Language, 71*(1), 1–43.

Bayne, T. (2010). *The unity of consciousness*. Oxford University Press.

Bayne, T., & Chalmers, D. (2003). What is the unity of consciousness? In A. Cleeremans (Ed.), *The Unity of Consciousness: Binding, Integration, Dissociation.* Oxford: Oxford University Press.

Block, N. (2007). Overflow, access, and attention. *Behavioral and Brain Sciences, 30*(5–6), 530–548.

Block, N. (2008). Consciousness and cognitive access. *Proceedings of the Aristotelian Society, 108*, 289–317.

Block, N. (2011). Perceptual consciousness overflows cognitive access. *Trends in Cognitive Sciences, 15*(12), 567–575.

Boly, M., et al. (2017). Are the neural correlates of consciousness in the front or in the back of the cerebral cortex? Clinical and neuroimaging evidence. *The Journal of Neuroscience, 37*(40), 9603–9613.

Bronfman, Z., Brezis, N., Jacobson, H., & Usher, M. (2014). We see more than we can report: 'Cost-free' color phenomenality outside focal attention. *Psychological Science, 25*, 1394–1403.

Brown, R. (2012). "The myth of phenomenal overflow." *Consciousness and Cognition, 21,* 599–604.

Chalmers, D. (1995). Facing up to the problem of consciousness. *Journal of Consciousness Studies, 2*(3), 200–219.

Chalmers, D. (1996). *The conscious mind: In search of a fundamental theory*. Oxford University Press.

Chalmers, D. (2015). Panpsychism and Panprotopsychism. In Y. Nagasawa & T. Alter (Eds.), *Consciousness in the physical world: Essays on Russellian monism.* Oxford: Oxford University Press.

Chalmers, D. (2017). The combination problem for panpsychism. In G. Brüntrup & L. Jaskolla (Eds.), *Panpsychism: Contemporary Perspectives.* Oxford: Oxford University Press.

Cohen, M., & Dennett, D. (2011). Consciousness cannot be separated from function. *Trends in Cognitive Sciences, 15*, 358–364.

Crane, T. (1988). The waterfall illusion. *Analysis, 48*(3), 142–147.

Dainton, B. (2000). *Stream of consciousness: Unity and continuity in conscious experience*. Routledge.

Dehaene, S., Changeux, J.-P., Naccache, L., Sackur, J., & Sergent, C. (2006). Conscious, preconscious, and subliminal processing: A testable taxonomy. *Trends in Cognitive Sciences, 10*, 204–211.

Dennett, Daniel. (1988). Quining Qualia. In A. Marcel & E. Bisiach (Eds.), *Consciousness in Modern Science.* Oxford University Press.

Dennett, D. (1991). *Consciousness explained*. Penguin Press.

Dennett, D. (1992). The self as centre of narrative gravity. In F. Kessel, P. Cole, & D. Johnson (Eds.), *Self and Consciousness: Multiple Perspectives.* Hillsdale, NJ: Erlbaum.

Duncan, J., & Humphreys, G. W. (1989). Visual search and stimulus similarity. *Psychological Review, 96*, 433–458.

Duncan, J., & Humphreys, G. (1992). Beyond the search surface: visual search and attentional engagement. *Journal of Experimental Psychology: Human Perception and Performance, 18*(2), 578–588. discussion 589–593.

Efron, R. (1967). The duration of the present. *Annals of the New York Academy of Sciences, 138*, 713–729.

Fink, S. (Ms.) 'How many stripes are on the tiger in my dreams?' https://philarchive.org/rec/FINHMS. Accessed October 2021.

Fink, S., Lyre, H., & Kob, L. (2021). A structural constraint on neural correlates of consciousness. *Philosophy and the Mind Sciences, 2*(7). https://doi.org/10.33735/phimisci.2021.79.

Freeman, J., & Simoncelli, E. P. (2011). Metamers of the ventral stream. *Nature Neuroscience, 14*(9), 1195–1201.

Geldard, F. (1982). Saltation in somesthesis. *Psychological Bulletin, 92*(1), 136–175.

Geldard, F., & Sherrick, C. (1972). The cutaneous 'rabbit': A perceptual illusion. *Science, 178*(4057), 178–179.

Goff, P. (2013). Orthodox property dualism + linguistic theory of vagueness = panpsychism. In R. Brown (Ed.), *Consciousness inside and out: Phenomenology, neuroscience, and the nature of experience.* New York: Springer.

Goff, P. (2017). *Consciousness and fundamental reality*. Oxford University Press.

Goff, P., & Roelofs, L. (Forthcoming). "In defence of phenomenal sharing." In J. Bugnon, M. Nida-Rümelin, & D. O'Conaill (Eds.), *The phenomenology of self-awareness and conscious subjects*. Routledge.

Gross, S., & Flombaum, J. (2017). Does perceptual consciousness overflow cognitive access? The challenge from probabilistic, hierarchical processes. *Mind and Language, 32*(3), 257–391.

Haun, A.M., Tononi, G., Koch, C., & Tsuchiya, N. (2017). Are we underestimating the richness of visual experience? *Neuroscience of Consciousness* (1)1–4.

Husserl, E. (1991). *On the phenomenology of the consciousness of internal time (1893–1917), trans. John Brough, Collected Works,* (Vol. 4). Kluwer Academic Publishers.

Kant, I. (1999). *Critique of pure reason (The Cambridge Edition of the Works of Immanuel Kant)*. Cambridge University Press.

Kolers, P., & von Grünau, M. (1976). Shape and color in apparent motion. *Vision Research, 16*(4), 329–335.

Kouider, S., de Gardelle, V., & Sackur, J. (2012). Do we still need phenomenal consciousness? *Trends in Cognitive Sciences, 16*, 140–141.

Kouider, S., & Dehaene, S. (2007). Levels of processing during non-conscious perception: A critical review of visual masking. *Philosophical Transactions of the Royal Society B, 362*, 857–875.

Lamme, V. (2015). The crack of dawn: Perceptual functions and neural mechanisms that mark the transition from unconscious processing to conscious vision. In T. Metzinger & J. M. Windt (Eds), *Open MIND: 22(T). Frankfurt am Main: MIND Group* (pp. 1-34). https://doi.org/10.15502/9783958570092.

Lau, H., & Rosenthal, D. (2011). Empirical support for higher-order theories of conscious awareness. *Trends in Cognitive Sciences, 15*(8), 365–373.

Levi, D. M. (2008). Crowding–An essential bottleneck for object recognition: A mini-review. *Vision Research, 48*, 635–654.

Mendelovici, Angela. (2020). Panpsychism's combination problem is a problem for everyone. In W. Seager (Ed.), *Routledge Handbook of Panpsychism*. New York: Routledge.

Miller, G. (2017). Can subjects be proper parts of subjects? The de-combination problem. *Ratio, 30*(2), 1–18.

Miller, G. (2018). Forming a positive concept of the phenomenal bonding relation for constitutive panpsychism. *Dialectica, 71*(4), 541–562.

Montero, B. (2017). What combination problem? In G. Brüntrup & L. Jaskolla (Eds.), *Panpsychism: Contemporary Perspectives*. Oxford: Oxford University Press.

Mørch, H. H. (2014). *'Panpsychism and causation: A new argument and a solution to the combination problem PhD dissertation*. University of Oslo.

Morrison, J. (2016). Perceptual confidence. *Analytic Philosophy, 57*(1), 15–48.

Nagel, T. (2004). 'Panpsychism' (originally published 1979). In D. S. Clarke (ed.), Panpsychism: Past and Recent Selected Readings. Albany: State University of New York Press

Oizumi, M., Albantakis, L., & Tononi, G. (2014). 'From the phenomenology to the mechanisms of consciousness: Integrated information theory 3.0.' *PLOS Computational Biology 10*(5), e1003588. https://doi.org/10.1371/journal.pcbi.1003588.

Pelli, D. G., & Tillman, K. A. (2008). The uncrowded window of object recognition. *Nature Neuroscience, 11*, 1129–1135.

Pockett, S. (2002). On subjective back-referral and how long it takes to become conscious of a stimulus: A reinterpretation of Libet's data. *Consciousness and Cognition, 8*, 45–61.

Pockett, S. (2003). How long is 'now'? Phenomenology and the specious present. *Phenomenology and the Cognitive Sciences, 2*, 55–68.

Roelofs, L. (2016). "The unity of consciousness, within and between subjects." *Philosophical Studies, 173*(12), 3199–3221.

Roelofs, L. (2019). *Combining minds: How to think about composite subjectivity*. New York: Oxford University Press.

Roelofs, L. (2020). "Can we sum subjects? Evaluating Panpsychism's hard problem." In W. Seager (Ed.), *The Routledge handbook of Panpsychism* (pp. 245–258). New York: Routledge.

Rosenthal, D. (2007). Phenomenological overflow and cognitive access. *Behavioral and Brain Sciences, 30*(5–6), 522–523.

Schwitzgebel, E. (2014). The crazyist metaphysics of mind. *Australasian Journal of Philosophy, 92*(2014), 665–682.

Seager, W. (1995). Consciousness, information and panpsychism. *Journal of Consciousness Studies, 2–3,* 272–288.

Seager, W. (2006). The 'Intrinsic Nature' argument for panpsychism. *Journal of Consciousness Studies, 13*(10–11), 129–145.

Seager, W. (2010). Panpsychism, aggregation and combinatorial infusion. *Mind and Matter, 8*(2), 167–184.

Seager, W. (2017). Panpsychist infusion. In G. Brüntrup & L. Jaskolla (Eds.), *Panpsychism: Contemporary Perspectives.* Oxford: Oxford University Press.

Schechter, E. (2013). Two unities of consciousness. *European Journal of Philosophy, 21*(2), 197–218.

Shoemaker, S. (2003). Consciousness and co-consciousness. In Axel Cleeremans (Ed.), *The Unity of Consciousness.* Oxford: Oxford University Press.

Sperling, G. (1960). The information available in brief visual presentations. *Psychological Monographs: General and Applied, 74*(11), 1–29.

Strawson, G. (2006). Realistic monism—Why physicalism entails panpsychism. *Journal of Consciousness Studies, 13*(10–11), 3–31.

Tononi, G. (2012). Integrated information theory of consciousness: An updated account. *Archives Italiennes De Biologie, 150*(2–3), 56–90.

Treisman, A., & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive Psychology, 12*(1), 97–136.

Treisman, A., & Sato, S. (1990). Conjunction search revisited. *Journal of Experimental Psychology: Human Perception and Performance, 16*(3), 459–478.

Tsuchiya, N., et al. (2015). No-Report paradigms: Extracting the true neural correlates of consciousness. *Trends in Cognitive Sciences. 19*(12), 757–770

Tye, M. (2009). A new look at the speckled hen. *Analysis, 69*(2), 258–263.

Tye, M. (2010). Up close with the speckled hen. *Analysis, 70*(2), 283–286.

Wertheimer, M. (1912). Experimentelle Studien über das Sehen von Bewegung [Experimental studies on motion vision]. *Zeitschrift Für Psychologie, 61*(1), 161–265.

Woodward, P. (2015). *'The emergence of mental content: An essay in the metaphysics of mind PhD dissertation*. Oslo: University of Oslo.

Woodward, P. (2021). "The selection problem for constitutive Panpsychism." *Australasian Journal of Philosophy, 99*(3), 564–578.

Zeki, S. (2007). A theory of micro-consciousness. In M. Velmans & S. Schneider (Eds.), *The Blackwell Companion to Consciousness.* Oxford: Blackwell.