

# The Limitations of Block's 'Overflow' Argument With Respect to the Possibility of the Study of Consciousness

S.E.R. CHERRY

St. Edmund Hall, University of Oxford

## I: THE PHENOMENAL/ACCESS DISTINCTION & THE STUDY OF CONSCIOUSNESS

Phenomenal consciousness [PC] and access consciousness [AC] are properties of mental states. After Nagel,<sup>1</sup> what makes a state P-conscious for Block is that 'there is something that it is like to be in that state'.<sup>2</sup> Accordingly, PC might be equated with experience, and the contents of PC at a given time with the totality of one's experiential states at that same time. As such, PC requires *Awareness*:<sup>3</sup> a mental state is *Aware* if and only if it its 'content is in some sense 'presented' to the self' or 'comes with a sense of ownership or... has 'meshness''.<sup>4</sup> By contrast a mental state is A-conscious if it can directly, rationally control our actions; AC does not require *Awareness* to have this property.

Whether PC and AC are in fact distinct is an empirical question, namely 'whether a subject can have an experience he does not and cannot think about',<sup>5</sup> or indeed vice-versa. But though Block presents this as a single question, I think there are two separate but related claims here:

### *Weak Claim*

A subject S can have a mental state which is P-conscious but which *isn't* A-conscious.

### *Strong Claim*

A subject P can have a mental state S which is P-conscious but which *can't* be A-conscious.

Clearly, the strong claim modally entails the weak one. If a participant can have at least one mental state which is P-conscious but which *can't* be A-conscious, then a participant can have a mental state which is P-conscious but which *isn't* A-conscious. But it does not follow that because any given P-conscious mental state isn't A-conscious that it couldn't have been made A-conscious.

This modal difference matters because if the strong claim is true, the possibility of there being a scientific study of consciousness is under threat. Presumably, to study and theorise

---

<sup>1</sup> Nagel, 1974

<sup>2</sup> Block, 2002: 206

<sup>3</sup> The term is capitalised because it's being used in a technical sense.

<sup>4</sup> Block, 2007

<sup>5</sup> Block, 2008: 289

about consciousness, we need to know at least superficially about the contents of conscious mental states. It seems reasonable to think of mental states in both AC and PC as being meaningfully ‘conscious’ according to our pre-theoretic understanding of consciousness. So any thorough investigation into or theory about consciousness should address the contents of both PC and AC.

How are we to investigate what the contents of an individual’s P-conscious mental states are? Reporting the content of one’s mental state (at least hypothetically) makes the mental state’s content indirectly accessible to a third-party.<sup>6</sup> And as it stands, reporting seems to be the only way of even superficially determining the contents of a participant’s P-conscious mental states.<sup>7</sup> But reporting the contents of a mental state requires it by definition to be A-conscious; for reporting is an action, and only contents of A-conscious mental states can directly influence our actions. So, currently, the only way for us to investigate the contents of a P-conscious mental state is for these contents also to be A-conscious.

The history of scientific progress warns any philosopher against boldly asserting that this will always be the case, i.e. that we will always be solely reliant on self-reporting to study the contents of a mental state. But until there’s a great methodological advance in the cognitive science of consciousness, we can only investigate the contents of P-conscious mental states which are also A-conscious. Hence, if there are mental states which are meaningfully conscious in the sense of being P-conscious, but which are inaccessible (i.e., if the strong claim is true), then we are likely limited in furthering our study of consciousness. For if we can’t gain knowledge about the contents of these mental states, our knowledge about the nature of such states of consciousness beyond the simple fact of their existence is limited. As Cohen & Dennett write in their criticism of Block (and others) theories of a PC/AC distinction: ‘inaccessible conscious states... inherently prevent the possibility of confirmation or falsification [of scientific hypotheses about the nature of conscious states]’.<sup>8</sup> Accordingly, any scientific investigation into or theorising about such conscious mental states is greatly restricted – at least, as long as we are reliant on reporting to know the content of mental states.

This problem is particularly relevant to the question of whether it’s even possible to give an answer to the so-called ‘hard problem of consciousness’ – namely, what it’s like to experience something, to be in a certain mental state.<sup>9</sup> If we can’t find out about the contents of some inaccessible P-conscious mental states even in the vaguest terms, then there is no hope *ab initio* of ever providing a rudimentary understanding of what it’s like to be in at least *some* mental states. And without such an understanding it isn’t at all clear how we could begin to provide a neuroscientific theory of the causes of the contents of such mental states (if this is even possible), or even to correlate mental state contents with neural states – for we would have no content to causally explain by, or associate with, neural states. At most a defender of physicalism could say that mental states are physical events, but they could never say *which* mental states are *which* mental events,<sup>10</sup> and so they could never provide an interesting or useful causal explanation of consciousness. So again, the existence of mental states in PC but not AC would greatly limit the possibility of a theory of consciousness.

---

<sup>6</sup> Chalmers, 2010

<sup>7</sup> Cohen & Dennett, 2011

<sup>8</sup> Ibid. 361

<sup>9</sup> Chalmers, 2010

<sup>10</sup> Nagel, 1974

But if no individual P-conscious mental state is inherently incapable of being A-conscious – if no P-conscious mental state is inherently inaccessible – then there must exist some conditions for any given P-conscious mental state by which it *can* be made A-conscious. Such conditions could be, for example, the use of a different cue in an attentional task. And as long as these conditions can be discovered for any mental state we wish to investigate then we can make any P-conscious mental state A-conscious, and thus make its contents susceptible to at least a superficial investigation and analysis. Thus if we can show that Block’s argument fails to support the strong claim, we can significantly reduce the threat his position poses to the possibility of deepening our understanding of consciousness. Of course, this is not to say that by undermining the support for one of Block’s particular claims that consciousness will magically become scientifically investigable – it’s merely to show that the cognitive scientist of consciousness has dodged this particular bullet.

## II: BLOCK’S ‘OVERFLOW’ ARGUMENT

As we have defined them, there is no reason *a priori* to think that the mental states in PC and AC must or mustn’t always coincide. To demonstrate that PC and AC are (at least weakly) distinct we must show empirically that at a given time AC and PC aren’t co-extensive: that there is some mental state (or information contained therein) which is in one but not the other. Block seeks to do this by claiming that there are perceptual contents of the visual system which are P-conscious but not A-conscious, meaning that the contents of PC is greater than or ‘overflows’ the contents of AC.<sup>11</sup>

It should be noted that Block does not explicitly argue for the strong claim on the basis of the overflow argument – though he does posit this as an explanation of behavioural findings from patients with visuospatial extinction.<sup>12</sup> Rather the target of this article are those like Cohen and Dennett who seem to suggest that those who buy into a PC/AC distinction (what they call ‘dissociative theories’) on the basis of overflow arguments are committing themselves to the strong claim, and thus thinking that there are – for the foreseeable future – significant limitations to the scientific study of consciousness.<sup>13</sup>

Block’s paradigmatic evidential basis for this overflowing comes from an experiment by Sperling.<sup>14</sup> Sperling displayed 3x4 arrays of alphanumeric characters to participants for 50ms, followed by a blank screen. Participants reported seeing all twelve stimuli in the array and identified them correctly as letters. However, when Sperling asked participants to spontaneously recall which letters they had seen where, they were unable to report the location and identity of more than four accurately. Yet when cued by a tone shortly after the array had disappeared to indicate which row of the stimulus to recall in what is known as the *partial report* condition, participants could almost always accurately report the location and identity of all the letters in whichever row was cued. This is known as the *partial report superiority effect*.

What Block interprets this effect to mean is that, at the point of array presentation, participants experience the entire array in all its detail, but are unable to access/report all the

---

<sup>11</sup> Block, 2011

<sup>12</sup> Block, 2008

<sup>13</sup> Cohen & Dennett, 2011

<sup>14</sup> Sperling, 1960

information they have. For they apparently *can* recall any bit of it in detail after presentation *if cued*. In which case, there must be a functional distinction between PC and AC. For there is information in sensory impressions that is experienced (and thus in PC) but which can't (globally) effect our actions – as demonstrated by it being unreportable – and thus must not be in AC.<sup>15</sup>

The crucial assumption in the argument here is the *counterfactual assumption*:

*Counterfactual Assumption*

‘Any aspect of experience present in a partial report condition *would* have been present even if some other partial report had been cued.’<sup>16</sup>

For it's this assumption which allows Block to claim that the whole array is present in PC, even if it doesn't make it into AC. This assumption in turn relies on the *independence* assumption – that ‘a subject's experience of the stimulus in a [partial report] condition is independent of which report is cued because the cue comes only after display offset’.<sup>17</sup> For otherwise the experience of the array and the experience of the cue could in fact be one, inseparable conscious event. In which case it would make no sense to talk counterfactually about what *would* have been recalled from a *given* presentation if another row had been cued. For a difference in cuing would result in a different experience altogether. While there may be good reasons to doubt the independence assumption,<sup>18</sup> as I will show later, it's Block's counterfactual assumption that prevents him from establishing the *strong* claim. And it's the strong claim that poses the greatest risk to the possibility of the study of consciousness.

Before proceeding we should distinguish between *generic* and *specific* phenomenology.<sup>19</sup> When participants in a Sperling-array task claim that they have seen ‘all’ of the array, they are able to report an experience of a generic phenomenon – namely that for every letter *c*, they see *c* as a letter-like form. But they can't report a *specific* phenomenon for every letter – that is, they can't report every letter *c* as being an ‘A’ or a ‘B’ or a ‘C’, etc.. That this sort of generic, perceptually degraded mental state is A-conscious for participants is not disputed. What Block must do for his argument to support either of his claims is first show that participants in a Sperling-array task have a P-conscious mental state of a *specific* phenomenon for each letter – even though demonstrably not all these mental states can be made A-conscious *on a given trial*. For this is what it would mean for PC to overflow AC.

If Block *is* able to show that participants in a Sperling-array task experience a specific phenomenon for each letter then this would appear to be sufficient (given the argument so far, together with the independence assumption) to support his weak claim – that a subject S can have a mental state which is P-conscious but which *isn't* A-conscious. But even if he can do this his argument so far on the basis of Sperling's experiment isn't sufficient to demonstrate

---

<sup>15</sup> Whether mental states which can affect some actions but not others count as being A-conscious is a matter of taxonomical debate. At the very least we might say they are not *fully* A-conscious. But for the sake of simplicity in this paper I will assume that for a mental state to be A-conscious it must be globally accessible to any decision/action process, and thus that unreportable mental states are necessarily not A-conscious.

<sup>16</sup> Phillips, 2011: 386 [emph. added]

<sup>17</sup> Ibid. 386

<sup>18</sup> Ibid.

<sup>19</sup> Ibid. 402

the stronger claim – that a subject S can have a mental state which is P-conscious but which *can't* be A-conscious. And it isn't clear how Block's argument ever could.

### III: THE COUNTERFACTUAL ASSUMPTION & THE POSSIBILITY OF THE STUDY OF CONSCIOUSNESS

Block's argument, as I have highlighted, relies on the consequence of the counterfactual assumption that (at least prior to the cue) any of the letters in the array *could* be recalled if they were subsequently cued. This assumption is necessary for him to argue evidentially that PC truly overflows AC. For it allows him to infer that, prior to the cue being given, a specific phenomenon for each letter exists in PC, and thus that it's only after the cue is given and some of these phenomena become A-conscious that others are incapable to be made A-conscious.

Without the counterfactual assumption, the only reason to think that PC overflows AC is on the basis of the self-reports of participants, who usually feel as if they experience a specific phenomenon for each letter in the array. But opponents of Block's approach can simply say that this experience is illusory – that in fact the participants only have a generic phenomenon for the unattended letters, even if they think they have a specific phenomenon for each letter. And the evidence, it seems to me, doesn't favour one interpretation over another.<sup>20</sup> It could be the case that people genuinely have P-conscious experiences of a specific phenomenon for each letter, and that these mental states overflows AC (and are thus unreportable). Or it could be the case that there are no A-conscious specific phenomena which weren't already P-conscious, and the purported P-conscious experience of a specific phenomenon for each letter is merely an illusion – the product of the existence of generic phenomena in both AC and PC for the unattended letters. Both of these contradictory hypotheses will produce the same outcomes in whatever tasks we give participants, as long as we are solely reliant on the reportability of the contents of a mental state to demonstrate that it's P-conscious. So Block's appeal to the self-reports of the participants is too speculative on its own to be good grounds to think that PC overflows AC.

So the counterfactual assumption is essential to motivating either of Block's claims. But doesn't it contradict the strong claim, that a subject P can have a mental state S which is P-conscious but which *can't* be A-conscious?

It depends on what it means to say that a mental state *can't* be made A-conscious. What Block can't mean, if he is to maintain the counterfactual assumption, is that there are mental states which *could never* be made A-conscious. As in, that there are P-conscious mental states for which there exist *no* conditions (say, a different attentional cue) by which they could ever be made A-conscious. Let's call this the *threatening version* of the strong claim, for it is this conclusion about the relationship between AC and PC which would threaten the possibility of studying consciousness. I hope it is clear that the threatening version of the strong claim is in direct contradiction with the counterfactual assumption.

For his argument to be coherent, what Block must therefore mean by the strong claim then is that, *after cueing*, there are P-conscious states which can't be made A-conscious. But this leaves open the possibility that there are conditions by which these states *could* have been made accessible – namely if they were cued. In which case these states aren't inaccessible and

---

<sup>20</sup> For a discussion of more recent empirical work on overflow and its failure to be persuasive, see Philips, 2016.

thus unreportable *simpliciter*, for there are conditions by which these states can be made reportable. So even if AC overflows PC, the way in which it does so poses no true threat for the possibility of the study of consciousness. For the contents of any mental state in PC *can* be made (at least superficially) reportable, if the correct condition obtains. (It would be an empirical task for cognitive science to determine which conditions were needed to make which P-conscious states A-conscious.)

So if someone in Block's shoes wanted to maintain the threatening version of strong claim then they would have to abandon the counterfactual assumption. But if they did this, it isn't clear how the argument could get off the ground in the first place, because there would be no way of even inferring what the contents of PC were (except those contents of PC which were also in AC) – not at least while the only indication of what is in PC is for individuals to report the contents of PC, which requires that these PC states are also A-conscious. But without any way of inferring what the contents of PC are over and above those PC mental states which are also in AC, there are no grounds for thinking that PC overflows AC – not at least if we think that the evidence from self-reporting isn't strong enough on its own. But if this is the case, then Block's argument fails at the first hurdle.

To summarise: Block's argument needs the counterfactual assumption to motivate the view that PC overflows AC. But if this counterfactual assumption is true, then there are no P-conscious mental states which can *never* be made A-conscious. In which case the threatening version of the strong claim must be false. And further rejecting the counterfactual assumption – as would be necessary if one wanted to make the threatening version of the strong claim – results in there being no reason to think in the first place that there are P-conscious mental states which aren't also A-conscious. In which case the argument for any version of the strong claim couldn't even get off the ground. So there is no good, coherent reason for believing the threatening version of the strong claim.

It should be noted that all this discussion leaves open whether AC could overflow PC – a direction Block has taken in other arguments, focussing on evidence from patients with hemispatial neglect.<sup>21</sup> Discussion of these matters is beyond the scope of the present paper, though they may prove more fruitful for Block and his adherents in demonstrating a factual distinction between PC and AC. But even if AC overflowed PC this wouldn't pose the same threat to the possibility of the study of consciousness that PC overflowing AC would. For it wouldn't mean that there are mental states which are meaningfully conscious yet unreportable.

#### CONCLUSION

Even if we accept Block's overflow argument for the division between AC and PC, it does not follow that there are P-conscious mental states which can *never* be made A-conscious. For Block's motivation for the existence of specific phenomena in PC relies on the counterfactual assumption that for any given trial any given phenomenon *could* be recalled, if there were different conditions with respect to (attentional) cueing. And this assumption stands in direct contradiction to the threatening version of the strong claim. Further, it's unclear how one could motivate the view that there really were specific phenomena in PC which could never be made A-conscious without the counterfactual assumption; not, at least, given the current impossibility of finding out about the contents of P-conscious mental states without reporting. But if this is the case, then there is no good reason to think that there are mental states in PC

---

<sup>21</sup> Block, 2008

whose contents, given the right conditions, are inaccessible and thus (at least superficially) unreportable.

In which case accepting Block's division between PC and AC on the basis of his 'overflow' argument doesn't require one to think that there are inherently unreportable, yet conscious, mental states – rather it implicitly requires one to *reject* the threatening version of strong claim. So *pace* those like Cohen and Dennett, one can comfortably maintain that there's a meaningful dissociation between AC and PC without necessarily threatening the possibility of the scientific study of consciousness.<sup>22</sup>

### REFERENCES

- Block, Ned. 2002. "Some Concepts of Consciousness." In *Philosophy of Mind: Classical and Contemporary Readings*, ed. David Chalmers, 206-218. Oxford: Oxford University Press.
- 2007. "Consciousness, Accessibility, and the Mesh between Psychology and Neuroscience." *Brain and Behavioral Sciences* 30 (5/6): 481-499.
- 2008. "Consciousness & Cognitive Access." *Proceedings of the Aristotelian Society* 108: 289-317.
- 2011. "Perceptual consciousness overflows cognitive access." *Trends in Cognitive Sciences* 15(12): 567-575.
- Chalmers, David J. 2010. *The Character of Consciousness*. Oxford: Oxford University Press.
- Cohen, Michael A., and Daniel C. Dennett. 2011. "Consciousness Cannot Be Separated From Function." *Trends in Cognitive Science* 15 (8): 358-364.
- Nagel, Thomas. 1974. "What Is It Like to Be a Bat?" *The Philosophical Review* 83 (4): 435-450.
- Phillips, Ian. 2011. "Perception and Iconic Memory: What Sperling Doesn't Show." *Mind and Language* 26 (4): 381-411.
- 2016. "No watershed for overflow: Recent work on the richness of consciousness." *Philosophical Psychology* 29 (2): 236-249.
- Sperling, George. 1960. "The information available in brief visual presentations." *Psychological Monographs: General and Applied* 74 (11): 1-29.

---

<sup>22</sup> The author extends his thanks to McIntosh, Noble and Le Poidevin for their advice while writing this article.