Phillips on Unconscious Perception and Overflow

Nicholas D’Aloisio-Montilla

Abstract
Phillips (Philosophy & Phenomenological Research, 93, 419–451, p. 433, 2016a) argues that Block faces a “serious internal challenge” in defending the claim that unconscious perception is of the same fundamental kind as conscious perception. This challenge is said to result from Block’s commitment to phenomenal overflow. However, in this paper, I demonstrate that Phillips’ (Mind & Language, 26, 381–341, 2011b) rejection of overflow likewise renders his view on unconscious perception “internally challenged” and therefore equally problematic.

Keywords Perception · Unconscious · Overflow · Attention · Access

1 Introduction
The question of whether unconscious perception (or “seeing”) is of the same fundamental kind as conscious perception (the "SFK" claim) has become a recent focus of debate in the philosophy of mind (Phillips 2016a, p. 419, Phillips and Block 2016; Block 2012, 2016; Berger and Nanay 2016). According to Burge (2010, p. 376), perception is a mental kind constituted by “objective, sensory representation by the individual”, and this requires the exercise of “capacities to represent environmental attributes, or environmental particulars, as the same, despite radically different proximal stimulations” (Burge, 2010, p.114). What is therefore clear is that perceiving a visual scene requires far more than just sensory registration (Phillips 2016a). If we assume representationalism as most modern cognitive scientists do, then perception is likely to require the construction of high-level representations that are personal and available to central agency (Burge 2010; Phillips and Block 2016). Given this stricter criterion, Phillips (2016a, Phillips and Block 2016) argues that the notion of unconscious perception is mistaken, and that “more empirical and theoretical work” is required before we can commit to a view on unconscious seeing capacities. By contrast, Block (2012, 2016, Phillips and Block 2016) has appealed to a body of research collected on clinical syndromes (such as blindsight or visual neglect) to illustrate that some atypical subjects appear to exhibit unconscious perceptual abilities.
A particularly interesting aspect of the discourse on unconscious perception is that most participants hold independent views on the capacity of phenomenal consciousness (what an experience is like). Block (1995, 2008, 2011, 2014) argues that the capacity of phenomenal consciousness overflows — that is, has a higher capacity than — what can be encoded in working memory under focal attention (i.e. accessed), whereas Phillips (2011a, b, 2016b) argues that such encoding is necessary for conscious experience. Moreover, Phillips (2016a, p. 433) argues that given Block’s commitment to overflow, his case for unconscious perception faces a “serious internal challenge”. In brief, on Block’s view it seems to be at least a conceptual possibility that subjects who report a lack of consciousness in completing perceptual tasks might, in fact, be conscious of items “without knowing it” (Phillips 2016a; Block 2011; Phillips in Phillips and Block 2016). This disconnect could be caused by “conservative responding” in which subjects are not as liberal in reporting their awareness of items when catch trials are used. Alternatively, lack of reporting could arise from a failure in the cognitive subsystems that detect conscious sensitivity (Phillips 2016a, b; Phillips in Phillips and Block 2016). To reject the possibility of unreported consciousness, one might propose a “subjective threshold for consciousness” so that direct reportability is necessary for phenomenology: without prior reportability, there could not have been any conscious experience (Phillips 2016a; Phillips in Phillips and Block 2016). This is perhaps similar to the “transitivity principle” in the case of (actualist) higher-order theories of consciousness, which implies that consciousness is necessarily paired with consciousness of a mental state. However, we then have an apparent contradiction with Block’s prior commitment to consciousness without access, since overflow necessarily implies that consciousness can occur without overt reportability. As Phillips summarizes (2016a, p. 433), “By Block’s own lights [on the rich capacity of consciousness] an interpretation of the studies which he cites which appeals to degraded conscious perception together with biased responding must be ‘on the table’.” And so, if overflow implies that unconscious seeing is just “degraded conscious perception” then “Block’s case for unconscious perception faces a very serious internal challenge” (Phillips, 2016a, p. 433).

However, this paper makes clear that Phillips likewise faces a “serious internal challenge” in his rejection of unconscious perception in the light of his own views on the overflow debate. It shows that Phillips is committed to one of the following disjuncts: (1) unconscious perception is of the same fundamental kind as conscious perception (which is inconsistent with his rejection of the SFK claim), or (2) consciousness has a higher capacity than cognitive access (which is inconsistent with his rejection of overflow). To arrive at this disjunction, we will briefly explore three experimental paradigms that overflow theorists cite as supporting evidence for their claims. The first two paradigms will be shown to be (at least, prima facie) compatible with the no-overflow account offered by Phillips and a rejection of unconscious seeing, but the timescale and demanding nature of the third, “high-resolution image paradigm” will illustrate an internal inconsistency in the viewpoint held by Phillips (Block 2011, p. 573; Sligte et al. 2008; Sligte et al. 2010). As a final point of introduction, this paper’s argument is not limited to demonstrating an internal challenge in Phillips’ positions on unconscious perception and

1 When representations are encoded into working memory, they are available to a range of consumer systems in the frontal workspace and so are accessed (Block 1995, 2011; Carruthers 2015; Dehaene et al. 2006). Phillips could, in principle, reject the claim that working memory encoding is constitutive for cognitive access, but his view still requires phenomenal consciousness to conjointly occur with an alternative form of access.
the overflow debate. Rather, it aims to demonstrate that given subjects' performance in the high-resolution image paradigm, any no-overflow view that appeals to the existence of unconscious performance abilities (as opposed to “generically” or “gist-like” conscious abilities) is committed to the existence of unconscious perception.

2 Sperling’s Paradigm

Support for the proposed distinction between (phenomenal) consciousness and cognitive access is found in an experimental paradigm first introduced by Sperling (1960). In what has now become an example of the “partial report paradigm”, subjects are momentarily presented with a stimulus of 12 letters. Since the stimulus is available to subjects for no more than a few 100 milliseconds, some sort of sensory buffer must be able to preserve the details of previously-presented letters. When subjects are then asked to freely recall the letters of the grid, they manage to correctly report just 4 or 5 items (Sperling 1960; Phillips 2011a; Block 2011). To investigate this limited reportability, Sperling cued subjects to a particular row of the grid after it had already been offset (i.e. removed from view). Remarkably, subjects managed to correctly report almost all letters of a row when that row was cued, even though cues could direct attention to any of the presented rows (Sperling 1960; Block 2011). This demonstrated that sensory memory encodes the identity-defining details of almost all 12 letters, even if only 4 are later reportable. Importantly, also, subjects insist that they “see” more than they can remember during the presentation of the stimulus (Block 2011, p. 567). This suggests that despite their limited recall, subjects possess a rich phenomenology of the grid’s letters. Thus, several philosophers have argued that the identities of all 12 letters are consciously represented by sensory memory, even though necessarily fewer items are accessed and reported via read-out from working memory (Block 1995, 2007, 2008, 2011; Bronfman et al. 2014; D’Aloisio-Montilla 2017, 2018).

Contrastingly, many opponents of overflow maintain that subjects encode the identities of the letters non-consciously until the cue brings a subset of them to the focus of attention (Brown 2014; Cohen and Dennett 2011; Phillips 2011a, b, Stazicker 2011; Dehaene et al. 2006; Baars 1988; cf. Block 2011)². They claim that the specific details of the letters cannot be included in conscious experience until they are brought under focal attention, and hence that conscious experience can emerge only after top-down attentional capture and working memory encoding. As Cohen and Dennett (2011, p. 359) summarize “participants can identify cued items because their identities are stored unconsciously until the cue brings them to the focus of attention”. This is perhaps one of the reasons why Dennett (1991, p. 302) is willing to also testament unconscious processing to blindsight patients on the basis that “verbal denials are supported by neurological evidence”, and so introspective reports, in this case, suggest the potential for unconscious perception. It therefore seems the standard interpretation of no-overflow is

² I have chosen to denote the status of the grid’s letters prior to the cue as “non-conscious” as this is the view taken by many no-overflow theorists. However, some no-overflow theorists (e.g. Kouider et al., 2010; Brown, 2014; cf. Block, 2011) have proposed that subjects have a generic (or gist-like) phenomenology of the grid’s letters prior to the deployment of focal attention.
committed to being supportive of the rich capacity of unconscious perception: subjects must necessarily generate “unconscious perceptual representations” of previously presented letters in the partial report paradigm\(^3\) (Phillips 2016b; Phillips in Phillips and Block 2016; Phillips 2016a). In unpacking the nature of the unconscious storage of letters in the sensory buffer, we will need to determine whether this implies that unconscious perception is of the same fundamental kind as conscious perception. In doing so, we should take care to avoid a potential conflation: although “evidence for non-conscious perception does not entail evidence for non-conscious WM [working memory]” (Stein et al. 2016, p. 2), the converse is not true, as evidence for unconscious working memory does entail evidence for unconscious perception.

Let us return to Burge’s framing of the perceptual kind, viz. that perception requires the construction of high-level representations by the individual (Burge 2010; Phillips 2016a; Phillips and Block 2016). It is clear that prior to the cue, the contents of sensory memory (both iconic memory and perhaps fragile VSTM, discussed below) contain the specific identities of the letters that are later cued and reported. Even if one argues against their (specific) phenomenal consciousness while the grid is visible to subjects, there must be “informational persistence” of the letters to enable subjects to perform as observed (Phillips 2011a, b, p. 404). Although Phillips (2011a, b) argues that phenomenal consciousness of the letters' identities is postdictively determined by the later arriving cue, he makes no specific claims on the nature of the informational persistence before the cue’s arrival. Rather, he argues that cueing subjects to a particular subset of items, up to a few hundred milliseconds after a visual stimulus is offset, retrospectively modifies what subjects are able to see consciously (Phillips 2011a; b). Moreover, iconic memory is likely to temporarily encode very detailed representations of the letters during this interval even if these representations are pre-categorical or constitute a “raw copy” of the presented stimulus. This is because the alphanumeric stimuli eventually require semantic processing (Keogh and Pearson 2014). We are therefore perhaps part of the way to demonstrating that the unconscious persistence of information prior to the cue partially satisfies Burge’s perceptual criterion for high-level representations, although more needs to be said about their attribution to the individual.

Answering the question of whether the iconic representations constructed prior to the cue reflect “personal level values and understanding” is far more challenging (Block in Phillips and Block 2016, p. 182; Phillips in Phillips and Block 2016; Phillips 2016a; Block 2016). It is clear that the representations automatically encoded in sensory memory are “accessible” in the sense that they could be cognitively accessed if focal attention were appropriately directed (Block, 2007; Block 2008; Block 2011; ). This implies that the identities of the letters are available to encoding in working memory or some other equivalent workspace module. Thus, we could perhaps form the argument that unconscious sensory representations are accessible to “central agency” prior to the cue (Block, 2007; Block 2008). This is the key point if we are to attribute

---

\(^3\) In a similar vein, some no-overflow proponents are supportive of the existence of “generic” phenomenology, such that uncued letters can, at best, reach a level of gist-like consciousness during the task (Kouider et al. 2010; see Block 2011 for review). However, we will not further explore this given that the target of this article is unconscious perception and space is limited.
these representations on the level of the individual. Burge (2010) has suggested that individual-level representations are those available to central agency, and so the sensory representations in Sperling’s paradigm might satisfy this criterion. In fact, any dispositional sense of access that these representations fall under is perhaps a stronger requirement than what is needed for unconscious perception as Phillips (Phillips and Block 2016, p.181) concedes: “My suggestion was only that when a representation is unavailable to central agency we lack a positive ground for attribution”. Thus, it does not matter whether the objective representation of letters is automatic because their availability to central agency can be potentially construed to evince individual-level processing (cf. Phillips in Phillips and Block (2016) on CFS).

However, the attribution of individual-level representations to iconic memory conflicts with what Phillips (2011b, p. 401) has argued in the past: “the [sensory memory] informational store is at the sub-personal level and that such representations occur in many cases where no conscious awareness correlates with them.” On what grounds is this determination made? One line of argument is to say that there can be no personal-level representations without attention (cf. Block 2008). Indeed, this is implied in (Phillips 2011a, p. 216) suggesting that attending to singleton features occurs on a personal level: “Rather, they are the potential results of attending to a location (or object or feature), which is something that occurs at the personal level.” However, a more obvious explanation for rejecting personal-level processing in the sensory buffer is that sensory memory representations cannot be attributed to the “whole person” because they, for the most part, occur without conscious awareness if attention is necessary for phenomenology (Phillips 2016a; Phillips 2011 cf. Block 2008). This is because on the no-overflow view, most content in the sensory buffer is not rendered conscious because focal attention is a limited resource. For the sake of argument, then, we will assume that in the case of Sperling’s paradigm, unconscious sensory memory representations can remain sub-personal until the presentation of the cue. On its presentation, Phillips (2011a, b) then suggests subjects that possess a personal, conscious perception of the cued letters, yet this is postdictively influenced such that no unconscious perception had previously occurred.

3 The Change Detection Paradigm

A more recent update on Sperling’s paradigm applies considerable pressure on Phillips’ rejection of the SFK claim (Landman et al. 2003; Sligte et al. 2008; Sligte et al. 2010). This section will show that the unconscious representations implicated by Phillips’ no-overflow view can be interpreted as satisfying at least two of the three criteria for perception, namely them being (1) high-level and, (2) available to central agency. In the (retro-cue) change detection paradigm, subjects are first presented with a “memory array” containing 8 items. Subjects are then presented with a “test array”

---

4 This is accentuated by the fact that Phillips supports an “externalist” postdictive account as opposed to a “Stalinesque” or “Orwellian” interpretation. The former need not posit an entirely unconscious status of letters prior to the cue’s presentation, whereas the latter two are likely to do so.
which remains identical to the memory array except that one item might or might not change orientation, size or identity. In the interval between the disappearance of the memory array and appearance of the test array, subjects are cued to an item that had originally appeared. Importantly, cues are withheld from the offset of the memory array by 1–4 s, and this delay is considerably longer than that in Sperling’s paradigm. The subjects’ task is to make a two-way, forced-choice discrimination on whether the cued item has changed. Subjects’ accuracy in detecting a change has been shown to be impressive even with the increased cue delay. As Block (2011, p. 571) notes, “with cue delays up to 4 seconds, capacities… are still well above the 4 items of working memory – 7 of 8 items at 1 second, 6 of 8 at 2.5 seconds and 5 of 8 at 4 seconds”. Given the extended timescale of cue delays in the task, Phillips cannot appeal to a postdictive explanation of subjects’ performance. As Phillips (2011b) acknowledges, postdictive effects tend to be seen in the order of hundreds of milliseconds following stimulus offset and so “it is far less plausible to account for the findings in terms of perceptual postdiction” (Phillips 2011b, p.403). For example, Thibault et al. (2016) showed that a delayed onset of attention could generate a delayed conscious experience, but this was eliminated after a 400 ms interval between the stimulus’ offset and the cue’s onset.

Phillips therefore takes a different tack in defending a no-overflow interpretation of subjects’ performance in the change detection paradigm. He argues that, pace Block’s account in which obtain a rich conscious experience of the memory array’s items prior to the cue, subjects instead rely on entirely “unconscious perceptual representations of the specific details of the first display which influence either the appearance of, or responses to, the test display without implying rich conscious experience of the memory display” (Phillips 2016b, p. 233; Phillips 2011a, b; cf. D’Aloisio-Montilla 2017; cf. Block 2011). In other words, it is claimed that subjects can perform successfully without ever consciously perceiving the item in which they report changes. This implies that the unconscious storage of most of the memory array’s items must last for up to 4 s during the interstimulus interval until the attentional capture of the cued item. This proposal has accordingly come under pressure, given that the only study which has tested for unconscious abilities during change detection showed subjects cannot perform to the same standard in masked conditions as they can in regular trials (Soto et al. 2011; Block 2011). I will choose not to contest whether unconscious working memory has the capability to support such a change, since it is not directly relevant to this paper’s inquiry (see D’Aloisio-Montilla 2017, 2018). Instead, I will explore whether Phillips requires unconscious perception to be of the same kind as conscious perception to explain such powerful unconscious recall in the change detection paradigm.

There is a clear difference in the unconscious informational persistence that is required to explain performance in the change detection paradigm compared with Sperling’s paradigm. In this paradigm, it must be the case that sensory memory content before and after the cue’s presentation remain sub-personal and low-level in nature. This is because on Phillips’ account there can be no (specific) phenomenal consciousness of the cued item at any point in time during the blank, interstimulus interval (see D’Aloisio-Montilla 2017 for a separate refutation of this claim). As Phillips (2011b, p. 407) notes, “we have no reason to deny that performance in Landman and Sligte’s change detection studies results simply from sub-personal, non-conscious informational persistence.” There are a number of
problems with the claim that an enduring unconscious persistence is not perceptual. First, a longer-lasting sensory memory store known as "fragile VSTM" (visual short-term memory) is responsible for maintaining the details of the previously presented items during the interstimulus interval. Unlike the transient iconic store, fragile VSTM is “based higher up in the visual system” and involves both perceptual categorization and high-level visual processing (Block 2011, p.574). For instance, the Kanizsa illusion has been shown to be already represented in fragile VSTM (Vadenbrouke et al. 2012; Phillips 2016b). Second, some content of the fragile VSTM representations are, a fortiori, available to central agency on the cue’s onset. Whatever it is that subjects utilize to make their discrimination must be encoded in working memory or a similarly related upstream consumer system. It therefore seems that some of the unconscious content that is held in fragile VSTM is necessarily placed into a position that is far closer to central agency than the dispositional accessibility of content in Sperling’s paradigm. Thus, it is clear that some details of the cued item are likely to enter “explicit memory”, unless we want to consider subjects as solely “guessing” at a potential change in the item, which would be at odds with recorded performance levels (Block 2011; Phillips 2011b). Thus, fragile VSTM representations—especially those that encode the details of the cued item—appear to satisfy two of Burge’s (2010) criteria for unconscious perception in their high-level nature and proximity to working memory encoding.

With regards to these representations being attributable to the individual, it is to be expected that most items represented in working memory can be influenced by beliefs and personal biases, given that working memory content is accepted to almost always be consciously experienced (Block 2011; Soto et al. 2011; Stein et al. 2016). This means that Phillips has to defend the claim that the cued representations which subjects utilize in making their judgement are not encoded in working memory – at least to the extent that a subject might be able to further conceptualize or reason about them. Prima facie, the way out for Phillips is to argue that the lack of complexity in making a two-way, forced-choice change discrimination renders it compatible with an exclusively sub-personal persistence of information. For instance, as he (2011b, p. 406) writes, in defence of the no-overflow explanation, “it is well established that people are worse at identifying items than detecting changes.” In fact, Phillips uses this line of argument to question whether blindseers at all require high-level processing in the successful completion of similar paradigms: “second, does successful forced-choice responding in blindsight reflect a direct visual response? A long-mooted alternative is that such responding reflects indirect awareness of autonomic reflex responses to visual stimulation” (Phillips 2016a, p. 435). The fundamental problem with an appeal to blindsight is that in the change detection paradigm, subjects are required to provide a considerably more complex judgement, namely the accurate reporting of a change in any one of 8 items offset for up to 4 s (Block 2011). Since subjects can be asked to determine a change in any one of many singleton features, this requires far more complex processing than is implicated in blindsight orientation paradigms (see Kentridge et al. 2004).

5 It is unclear whether or not “full-blown” cognitive access ensues, given many theorists take cognitive access to require phenomenal consciousness (e.g Baars 1988; Carruthers 2015).
It therefore seems that Phillips’ rejection of the SFK claim initially faces an internal challenge from his explanation of performance in the change detection paradigm. Prima facie, it is unclear why subjects do not rely on unconscious perceptual representations if the overflow view is to be rejected, and is replaced with an account that posits entirely unconscious persistence of the cued item. On Phillips account, these unconscious representations could be naturally interpreted as perceptual, in the sense that they appear to be high-level fragile VSTM representations that can be made available to central agency, and are potentially attributable to the individual. Given the latter claim on the individual-level capacities of sensory memory is the most controversial, the next section cites a third paradigm to make clear that some of the unconscious persistence of fragile VSTM content should be interpreted as perceptual on Phillips’ no-overflow account.

4 The “High-Resolution Imagery Paradigm” (Block 2011, p. 573)

I will now demonstrate that Phillips’ no-overflow view seemingly allows for unconscious persistence that is further attributable to the individual — in addition to being high-level and available to central agency — thereby satisfying Burge’s (2010) framing of unconscious perception. The task demands of the change detection paradigm were extended by Sligte et al. (2010). In this variant task, subjects were required to report the original identity of the cued item if it had changed, in addition to registering the two-way, forced-choice discrimination. To do this, subjects were presented with four items at the end of change trials, and were asked to select the item which was no longer present in the test array. Importantly, subjects were cued to an item after a 1-s delay from the memory array’s offset, thus again ruling out a sole reliance on early iconic memory or a postdictive explanation of subjects’ performance. Sligte et al. (2010) reported that subjects’ performance in both the change detection and change identification tasks were successful, with subjects managing to correctly identify the changed item 80.6% of the time when coloured items were presented, and 76.7% of the time with greyscale items. This indicates that subjects can complete both a four-way and two-way forced-choice discrimination by relying on high-level representations that are maintained in fragile VSTM. Overflow theorists argue that these fragile VSTM representations were rendered phenomenally conscious with a conscious icon “outside of focal attention” (Block 2011, p. 569). However, given Phillips’ (2011a, b) position on the overflow debate, he is committed to suggesting that performance in both discriminations can occur without conscious recall.

Although Phillips has yet to make reference to this paradigm, his no-overflow interpretation must again be committed to the idea that not only reporting the presence of a change, but further reporting the original identity of previously presented items of the memory array is “compatible with a lack of conscious recall of the initial memory array-array” (Phillips 2011b, p. 406). Thus, to demonstrate that this violates Phillips’ position against unconscious perception, I will show that unconscious fragile VSTM representations used in the high-resolution imagery paradigm must be closely linked to individual-level processing, since it is clear from the regular change detection paradigm that these representations are similarly high-level and available to central agency. I will therefore make the case that subjects must either recruit unconscious perception in
completing the task, or alternatively, make use of conscious representations thereby contradicting Phillips’ position on overflow. Sligte et al. (2010, p. 2) have labelled the representations that subjects rely on in the four-way change identification task as “high-resolution object representations”. Given the fact that subjects have to compare and identify the cued item against one of four candidates, this suggests that subjects make use of a range of features “bound” into a single object representation, as opposed to free-floating singleton features (Sligte et al. 2010, p. 9). As Sligte et al. (2008, 2010, p. 7) argue, “As change detection and identification more often than not will depend on different feature dimensions, a higher degree of feature binding would be necessary to perform correctly on change detection and change identification…”. Moreover, in certain identification display configurations, all four items might have overlapped in at least one visual dimension (such as the color red), and this rules out performance that is solely based on the recall of unbound singleton features (Sligte et al. 2010). Thus, the unconscious object representations implicated in this paradigm by Phillips are far more impressive in their necessary (1) binding, (2) detail, and (3) high-level content when compared with those recruited in the two-way, change detection paradigm.

This means that in order for Phillips to uphold his rejection of the SFK claim, unconscious object representations must be attributed to a particular subsystem instead of the whole person or a central module such as working memory. However, I have seen no suggestion in the literature that bound “object representations” – as opposed to the encoding of singleton features or fragments – can be easily maintained at the sub-personal level. For instance, object representations are known to elicit the same-object advantage, yet there have been far fewer studies to demonstrate that object-based attention can act on unconscious stimuli than, say, feature-based or spatially-based attention (see Norman et al. 2013). This further suggests that object representations are closely linked to consciousness and thus personal-level processing. Furthermore, Phillips (2016a, p. 424) claims that cases of unconscious processing in neglect might be explained by “non-perceptual sensory registration of…features”. However, the object representations in the change identification task encode the details of items such as cars, butterflies or traffic lights which contain complexly mapped features that are considerably more “abstract” in nature (Block 2011; Sligte et al. 2010). There is clearly a lack of evidence to suggest that unconscious representations with such “high-resolution” are solely confined to sub-personal systems during the identity report, even if they are to remain outside of explicit conscious access (in Phillips and Block 2016). In other words, there is a strong reason to suggest that these unconscious representations implicate individual-level processing.

Moreover, if one’s ability to unconsciously recall the actual identity of previously presented items (that are encoded in object representations) only amounts to the “pre-perceptual registration of features”, then it seems that Burge’s definition of perception might be called into question (Phillips and Block 2016; Sligte et al. 2008; Sligte et al. 2010). This is because currently known unconscious abilities, tend to be only able to make correct judgements in two-way (and not four-way) discrimination tasks—for instance, such as whether a currently available item is horizontal or vertical in nature (Kentridge et al. 2004; Block 2011). By contrast, significantly more processing is required in the high-resolution image paradigm, as subjects are
required to report both a change and identity in any one of eight items multiple seconds after a stimulus has been offset. Thus, if we assume that subjects can complete this task unconsciously as Phillips (2011b) suggests, then the processes required to complete such a demanding change discrimination seem to implicate perceptual processing in a way that blindsight need not. Recall that object representations are closely linked to availability to working memory encoding, which itself is characteristically susceptible to influence from personal-level judgements, beliefs and desires. If it is clear that in order for one to unconsciously report the identity of a previously presented item in the high-resolution image paradigm, both object representations and some availability to central agency is necessary, then we should not accept a necessary sub-personal persistence of the reported item without substantial empirical support. In fact, there is a strong reason to suggest that unconscious identification of items can be influenced by personal-level beliefs and processes in a way that is absent in the two-way forced-choice change discrimination. Prima facie, reporting the identity of a previously presented item requires far more work than noticing a change in that item. Neither Phillips (2011a, 2016a, b; Phillips and Block 2016) nor Block (2011, 2016) have suggested that the former can occur without perception. Take for example, choosing whether a butterfly or flower had been previously presented in the memory array. A critic of Phillips could, in turn, argue that unconscious perceptual representations of the stimulus interact with belief systems, perhaps linked to object identification, and so the eventual item report could be biased by a preference for one of the two items (Block in Phillips and Block 2016).

The strongest argument, however, for the requirement of unconscious perception in Phillips’ no-overflow account arises from questioning whether the cued item can unconsciously persist without individual-level processing. If we assume, for the sake of argument, that most items of the memory array are not “perceived” in Burge’s sense, then it still does not follow that the cued item is unperceived. The cued item is the subject of further attentional amplification and refinement once top-down attention is directed to the underlying object representations held in fragile VSTM. Since this implicates a further role for working memory, it becomes yet more difficult to argue that such intentional or “voluntary” behaviour of the subject (attention is mostly personal) towards bound, object representations does not require processing above the sub-personal threshold. This starkly contrasts what can be said of Sperling’s paradigm and the postdictive account of performance. It is worth noting that Block (2011) and overflow theorists take the performance of subjects in the high-resolution image paradigm to be impressive enough that it requires conscious perception without cognitive access at some point during the task (D’Aloisio-Montilla 2017). However, Phillips must object to any perception – either conscious or unconscious – as being necessary to complete the paradigm, since the former would contradict his no-overflow position and the latter would stand against his rejection of the SFK claim (Phillips 2016a, b). Although the high-resolution image paradigm implicates object representations that are far richer than what are required in the two-way forced-choice change detection paradigm, Phillips still requires subjects to rely on an implicit and sub-personal persistence of information. On his view, we would therefore arrive at the conclusion that that subjects
are capable of reporting the identities of previously presented items without unconsciously perceiving them or consciously “seeing” them, which would be (at least, prima facie) an undesirable outcome for Burge’s notion of perception.

Importantly, empirical evidence stacks up against both of these claims (Block 2011; Soto et al., 2011; Block 2016; Block in Phillips and Block 2016; Kouider et al. 2010). Through the use of visual masking, Soto et al. (2011) prevented subjects from consciously representing a memory array during a change detection paradigm. This meant that subjects relied on unconscious representations in order to complete their verbal report. As a result, we would expect subjects to perform well in this much simpler task if they can complete regular change detection without conscious experience. However, performance in masked trials was poor (at a slightly above chance rate), demonstrating the limited capacity of an unconscious strategy. In fact, subjects managed to only correctly detected changes at a slightly above chance rate of 55% (Soto et al. 2011). This implies, contra Phillips, that subjects might in fact be required to consciously perceive the item which they correctly identify in the high-resolution image paradigm. With this in mind, then, there is significantly less motivation to posit that subjects can solely rely on unconscious perception in the high resolution image paradigm, and even less reason to support the view that subjects can rely on a weaker memory persistence that is unconscious but not perceptual. Unconscious perception still enables subjects to rely on the same fundamental kind of perceptual processing that they might otherwise rely on in conscious tasks, whereas without unconscious perception, subjects must rely on a type of sub-personal processing that appears to have “been dislodged from by powerful evidence” (Block 2011, p.575). Thus, it seems highly likely that any unconscious processing must be perceptual if it is to solely account for subjects being able to accurately report the identity of a changed item in the high-resolution image paradigm.

Given that this evidence collected on the limited capacity of unconscious memory independently calls into question whether subjects can complete the four-way forced choice discrimination task without (specifically) conscious representation (as overflow theorists posit), then it is considerably more difficult to hold the additional position that subjects complete the task without some type of perception—unconscious or otherwise (Sligte et al. 2010; Block 2011; D’Aloisio-Montilla 2017; Vandenbroucke et al. 2011). Unconscious perceptual representations necessarily entail a lower threshold for neural activation than conscious perceptual representations. Following this reasoning, then, the baseline in activity required for unconscious persistence that is not perceptual is less than what is required for unconscious perception. In other words, the activation required to support unconscious, non-perceptual activity is significantly less than what is required in unconscious perceptual activity or conscious perception. Thus, it would surprising to find, given the contemporary evidence refutes the existence of rich unconscious representation, that subjects effectively employ unconscious representations that require the lowest activation signal of the three possible types of informational persistence when reporting the original identity of a changed item. Such a position seems incompatible the contemporary evidence on both consciousness and perception with regard to change detection paradigms, and further leads to the conclusion that subjects can report the identity of an item (not just a change in its properties) without “seeing” or “perceiving” it – conflicting our current
intuitions on the subject. Do we really want it to be possible for a subject to report the identity of a presented item – and thus a range of conjoined properties associated with it – without requiring them to at least unconsciously perceive it? Thus, in the interests of preferring a hypothesis that is better supported by empirical evidence collected on unconscious memory and unconscious reporting, it seems that if we are to accept the no-overflow position of Phillips, then we should be committed to the richness of unconscious perception (i.e. the SFK claim) to explain performance in this paradigm (Block 2011, p. 575):

As Phillips puts it, ‘[m]y claim is only this: for all that has been said, we have no reason to deny that performance in Landman and Sligte’s change detection studies results simply from sub-personal, non-conscious informational persistence’ ([31], p.407). If instead we adopt the methodology of asking which hypothesis is better supported, I think we should prefer the overflow hypothesis.

5 Concluding Remarks

Of course, a lot more can be said. However, this paper has demonstrated that Phillips’ rejection of the SFK claim—that is, unconscious perception being of the same fundamental kind as conscious perception—is seemingly inconsistent with his position in the overflow debate. Phillips’ no-overflow position is unable to explain performance in the high-resolution image paradigm without further implicating the existence of unconscious perception that is “objective, sensory representation by the individual” (Burge 2010, p. 376; Phillips and Block 2016). This is because that paradigm implicates the existence of far richer unconscious perceptual representations than what is required in two-way, forced-choice change detection paradigms. A few open questions remain. First, can adopting generic phenomenology save Phillips’ no-overflow interpretation while discounting any need for unconscious seeing? Second, can the high-resolution image paradigm be modified to demonstrate that representations required in reporting the identity of a changed item require individual-level processing capacities (for example, do personal beliefs or biases impact the accuracy of discriminating certain objects)? In any case, what is clear is that Phillips and Block face equally problematic “internal challenges” in the unconscious perception debate.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

I’d like to thank Peter Millican, Patrick Butlin and those close to me for comments on earlier versions of this manuscript.
References

Baars, B. (1988). *A cognitive theory of consciousness*. Cambridge: Cambridge University Press.

Baars, B. (2002). The conscious access hypothesis: Origins and recent evidence. *Trends in Cognitive Sciences*, 6, 47–52.

Baars, B., et al. (2013). Global workspace dynamics: Cortical binding and propagation enables conscious contents. *Frontiers in Psychology*, 4, 200.

Bartolomeo, P. (2008). The neural correlates of visual mental imagery: an ongoing debate. *Cortex*, 44, 107–108.

Berger, J., & Nanay, B. (2016). Relationalism and unconscious perception. *Analysis*, 76(4), 426–433.

Block, N. (1995). On a confusion about a function of consciousness. *Behavioral and Brain Sciences*, 18, 227–287.

Block, N. (2007). Consciousness, accessibility and the mesh between psychology and neuroscience. *Behavioral and Brain Sciences*, 30, 481–548.

Block, N. (2008). Consciousness, accessibility and the mesh between psychology and neuroscience. *Behavioral and Brain Sciences*, 30, 481–548.

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

Block, N. (2012). The grain of vision and the grain of attention. *Thought: A Journal of Philosophy*, 1(3), 170–184.

Block, N (2014). Rich Conscious Attention Outside Focal Attention, Trends in Cognitive Sciences, 18, 445–447.

Bronfman, Z. et al. (2014) We see more than we can report: ‘cost free’ color phenomenality outside focal attention. *Psychological Science*, 25, 1394–1403.

Brown, R (2014). Consciousness doesn’t overflow cognition. *Frontiers in Psychology*, 15, 1399.

Block, N. (2016). The Anna Karenina Principle and Skepticism about unconscious perception. *Philosophy and Phenomenological Research*, 93, 452–459.

Burge, T. (2010). *The origins of objectivity*. Oxford: Oxford University Press.

Carruthers, P. (2015). *Block: The origins of objectivity*. Cambridge: Cambridge University Press.

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

Cohen, M. A., Cavanagh, P., Chun, M. M., & Nakayama, K. (2012). The attentional requirements of consciousness. *Trends in Cognitive Sciences*, 16, 411–417.

D’Aloisio-Montilla, N. (2017). Imagery and Overflow. See see more than we report. *Philosophical Psychology*, 30(5), 545–570.

D’Aloisio-Montilla, N. (2018). A Brief Argument For Consciousness Without Access. *Ratio*, 31(2), 119–136.

Dehaene, S. (2014). Consciousness and the brain: Deciphering how the brain codes our thoughts. New York: Viking Press.

Dehaene, S., et al. (1998). A neuronal model of a global workspace in effortful cognitive tasks. *Proceedings of the National Academy of Sciences of the United States of America*, 95, 14529–14534.

Dehaene, S., Changeux, JP., Naccache, L., Sackur, J., Sergent, C. (2006) Conscious, preconscious, and subliminal processing: A testable taxonomy. *Trends in Cognitive Sciences* 10(5), 204–11.

Dennett, D. (1991). *Consciousness explained*. London: Allen Lane, The Penguin Press.

Dennett, D. (2001). Are we explaining consciousness yet? *Cognition*, 79, 221–237.

Kentridge, R. W., et al. (2004). Spatial attention speeds discrimination without awareness in blindsight. *Neuropsychologia*, 42, 831–835.

Keogh, R., & Pearson, J. (2014). The sensory strength of voluntary visual imagery predicts visual working memory capacity. *Journal of Visualization*, 14(12), 7, 1–13. https://doi.org/10.1167/14.12.7. Retrieved from http://jov.arvojournals.org/article.aspx?articleid=2193915.

Kouider, S., de Gardelle, V., Sackur, J., & Dupoux, E. (2010). How rich is consciousness? The partial awareness hypothesis. *Trends in Cognitive Science*, 14, 301–307.

Landman, R., Spekreijse, H., & Lamme, V. A. F. (2003). Large capacity storage of integrated objects before change blindness. *Vision Research*, 43, 149–164.

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

Norman, Liam J. and Heywood, Charles A. and Kentridge, Robert W. (2013) Object-based attention without awareness. *Psychological science.*, 24(6), 836–843.

Phillips, I. B. (2011a). Attention and iconic memory. In C. Mole, D. Smithies, & W. Wu (Eds.), *Attention: Philosophical and psychological essays* (pp. 204–227). Oxford: Oxford University Press.

Phillips, I. B. (2011b). Perception and iconic memory. *Mind & Language*, 26, 381–341.
Phillips, I. (2016a). Consciousness and criterion: on Block’s case for unconscious seeing. *Philosophy & Phenomenological Research, 93*(2), 419–451.

Phillips, I. B. (2016b). No watershed for overflow: Recent work on the richness of consciousness. *Philosophical Psychology, 29*(2), 236–249.

Phillips I, Block N (2016) Debate on unconscious perception. In: Nanay B (ed.), *Current Controversies in Philosophy of Perception*. New York: Routledge, 165–92.

Rerko, L., Souza, A. S., & Oberauer, K. (2014). Retro-cue benefits in working memory without sustained focal attention. *Memory and Cognition, 42*, 712–728. https://doi.org/10.3758/s13421-013-0392-8.

Sligte, I. G., et al. (2008). Are there multiple visual short term memory stores? *PLoS One*, 3, 1–9.

Sligte, I.G., et al. (2010). Detailed sensory memory, sloppy working memory. *Front. Psychol.*, 1, 1–10.

Soto, D., Mantyla, T., & Silvanto, J. (2011). Working memory without consciousness. *Current Biology, 21*, R912–3.

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

Stazicker, J. (2011). Attention, visual consciousness and indeterminacy. *Mind & Language, 26*, 156–184.

Stein, T., Kaiser, D., & Hesselmann, G. (2016). Can working memory be non-conscious? *Neuroscience of Consciousness, 6*, niv011.

Thibault, L., van den Berg, R., Cavanagh, P., & Sergent, C. (2016). Retrospective attention gates discrete conscious access to past sensory stimuli. *PLoS One*, 11(2).

Tye, M. (1995). *Ten problems of consciousness*. Cambridge: MIT Press.

Vandenbroucke, A. R. E., et al. (2011). Manipulations of attention dissociate fragile visual short term memory from visual working memory. *Neuropsychologia, 49*, 1559–1568.

Vandenbroucke, A. R. E., Sligte, I. G., Fahrenfort, J. J., Ambroziak, K. B., & Lamme, V. A. F. (2012). Non-attended representations are perceptual rather than unconscious in nature. *PLoS ONE, 7*, e50042.