Does Phenomenal Consciousness Overflow Attention? An Argument from Feature-Integration

Winner of the Gerritt and Edith Schipper Undergraduate Award for Outstanding Undergraduate Paper at the 62nd Annual Meeting of the Florida Philosophical Association

Joshua Myers, *University of Miami*

Introduction

Are you phenomenally conscious of the sensation of your clothes against your skin? Surely you are now, but were you before you read the question? Or, did you only begin to experience the tactile sensation of your clothes once you directed attention to it? People differ widely in their intuitions on the matter (Schwitzgebel 2011). Regardless, this simple exercise goes a long way towards motivating and bringing to light a rough, first approximation of the question this paper will be concerned with: Is attention necessary for phenomenal consciousness? Cognitive scientists have discovered phenomena such as inattentional blindness and change blindness in which subjects fail to report even glaringly incongruous objects that are not attended to, suggesting that their reported lack of phenomenology is caused by their inattention (Simons & Levin 1997). On the other hand, some experiments, along with some of our strongest intuitions, suggest that we may be conscious of much more than we direct attention to, which is what Ned Block calls phenomenal overflow (Block 2007). In this paper, I will ask whether attention is necessary for consciousness, and after weighing the evidence, I answer in the affirmative.

First, let me define my terms. Ned Block first introduced "access consciousness" in opposition to the more familiar, phenomenal, usage of "consciousness" (Block 1995). Phenomenal consciousness is the subjective, experiential, "what-it-is-like" quality of an experience while access consciousness is the voluntary and rational (as opposed to automatic) use of information in cognition. The question then becomes whether a dissociation can be drawn between these two types of consciousness. In particular, I will focus my discussion on whether it is possible for phenomenal consciousness to be present in the absence of access consciousness. There is some confusion about whether access consciousness should be understood dispositionally or categorically. In other words, does information merely have to be *accessible* or does it have to be *accessed* in order to be access conscious (Carruthers 2015)? Following Block (2011), I will assume the latter, stronger, interpretation. This is because I take it to be a more philosophically interesting question whether or not representations that are not

accessed (but may or may not be accessible) at time t_1 can still be phenomenally conscious at time t_1. This is the reading Block gives when he argues that simply because representations "are not accessed is not to say that any are inaccessible" (Block 2011). So, phenomenal overflow is a dissociation of phenomenal consciousness from access, not accessibility.

I will focus my discussion on one particular access mechanism that has received the most interest in the literature: attention.[1] On the standard account, attention is a kind of filter that allows important perceptual information to be accessed, or made available for global processing. That is to say, attention is trivially linked to access consciousness. Attention can be bottom-up, as when it is automatically drawn towards a salient object in a visual scene, or top-down, as when it is consciously and intentionally directed towards an object or area of a visual scene (Srinivasan et al. 2009). The attentional "filter" allows us to focus our processing on the contextually relevant stimuli and typically results in enhanced information processing and better performance on a given task (Srinivasan et al. 2009).

For the remainder of this paper I will use "phenomenal consciousness overflowing attention" interchangeably with "phenomenology overflowing access." The line of argument I pursue in the paper is meant to directly answer Ned Block's arguments for overflow. However, the latter terminology, on Block's view, technically implies a broader category of phenomena than the former since attention is merely one of many access mechanisms. The only access mechanism that this paper will be concerned with is *attentional* access. At this point, before my argument has even gotten off the ground, Block might object that we are talking past each other. If Block is concerned with a plurality of access mechanisms and my argument is only directed at one of them then I have will not have refuted Block's argument for overflow. However, I think this focus on attention can be motivated for two reasons. Firstly, as we shall see, all of Block's positive arguments for overflow, and most notably the Sperling (1960) paradigm, hinge on some form of visual attention. Thus, in the absence of reasons to accept that Block's arguments generalize to other access mechanisms, by showing that phenomenal consciousness does not overflow attention I will have undercut the motivation for Block's view.[2] Secondly, the literature on overflow has converged on attention as the most interesting access mechanism in the visual modality that could plausibly be implicated in the kind of global broadcasting and availability that is the hallmark of access consciousness (Brigard & Prinz 2010, Cohen & Dennett 2011, Woodman & Luck 2003). Thus, in the context of arguments concerning phenomenal overflow in the visual modality, and specifically the arguments put forward by Block (2007, 2011), attention is the access mechanism in question. So, although my arguments do not show that phenomenal overflow *never* occurs, they do undercut the main motivation for positing overflow in the first place. In other words, I will argue against Block *by way of* showing that phenomenal consciousness is commensurate with attention.

There are a number of methodological, epistemic, and theoretical implications that the relationship between phenomenal consciousness and attention might have for philosophy and cognitive science. For example, it may be important to the methodology of how consciousness is studied. If some phenomenally conscious states can be inaccessible or otherwise unaccessed, then traditional methodologies employed in cognitive science such as self-report and voluntary action may be bypassing or otherwise overlooking certain conscious states of interest (Cohen & Dennett 2011). The very fact that there is a debate over phenomenal overflow may also call into question the reliability of introspection. After all, either our introspective intuition of rich phenomenology is incorrect *or* we have conscious representations that cannot be accessed by introspection alone. It is surprising that such radically different hypotheses about what we consciously experience are seemingly irresolvable by introspection alone (Schwitzgebel 2011). Another epistemic upshot of the relationship between phenomenal consciousness and attention is the role attention might play in justifying our beliefs. Given the intuitive plausibility of the role of phenomenally conscious experiences in justifying perceptual beliefs, then attention may or may not be a necessary condition for perceptual justification (Siegel & Silins 2014). Perhaps most centrally, the relationship between consciousness and attention clearly has major implications for the cognitive and neural architecture of consciousness. Theories of consciousness such as the higher-order thought theory and global workspace theory, both of which posit that unaccessed representations cannot be conscious, will stand or fall with the question of overflow (Dretske 2007, Brigard & Prinz 2010, Carruthers 2015). If mental states can be phenomenally conscious independent of their being accessed, then phenomenal consciousness is a first-order property of certain mental states that does not depend on any higher-order states or global availability.

In this paper, I will argue that we do not have good reasons to accept arguments for the possibility of consciousness without attention and that, in fact, attention is necessary for consciousness. In order to do this, I will first present the strongest arguments for consciousness overflowing attention. Then I will show that there are tenable alternatives to these arguments that explain the same data without positing unattended conscious states, thereby undercutting any reason to accept overflow in the absence of further arguments. The key to this argument is that unconscious visual processing possesses the same explanatory power as conscious visual processing in interpreting that data. Next, I will develop one of these alternatives to overflow that accounts for our intuition of phenomenal richness by appealing to the feature-integration theory of attention.

Arguments for Overflow

Ned Block, one of the staunchest advocates of overflow, draws much of his empirical support from a 1960 experiment by George Sperling (Block 2007, Block 2011). In it, subjects were shown a 3x4 matrix of letters for 50 milliseconds (ms). In control experiments subjects could only report up to four letters, or a third of the matrix. The reported letters were randomly distributed throughout the matrix. On the experimental trials a tone was played after the stimulus disappeared that indicated to the subjects which row of the matrix they had to report. The tone was played after the visual display disappeared so as not to draw attention to any particular area of the matrix beforehand. This is known as a retro-cue. Surprisingly, subjects were able to report on average all four symbols in the signaled row after the retro-cue. The fact that the row was retro-cued makes these results significant, for the matrix was not present for the subjects to consult and attention could not be directed to the cued row beforehand. These results did not change as a function of how long the stimulus was presented, as long as it was kept between .015 and .500 seconds (Sperling 1960). This experiment has been replicated with 32 alphanumeric characters (Sligte et al. 2008).

The object limit for working memory, as determined by the non-cued results, is approximately four (Block 2007).[3] However, subjects' ability to report the entire cued row without directing attention to it until *after* the display disappeared is initially inexplicable. Since subjects were not able to simply look at the display and report what they saw after the cue, and since the control trials indicated that subjects were only able to remember four random symbols out of the whole display, Block argues that subjects must have had a phenomenally conscious representation of the whole display in iconic memory that was being consulted in order to report the entire row (Block 2007). If this is true, then this paradigm provides a clear example of consciousness (of the whole matrix) overflowing attentional access (of merely the single row of letters).[4] Crucially, this argument relies on the fact that subjects "insist that they have seen more than they can remember afterwards" which seems to indicate that there was conscious phenomenology that could not be accessed (Sperling 1960). Notice that there is a trivial way of interpreting the Sperling paradigm that posits that subjects saw the entire matrix but simply forgot it. We forget things all the time; why is this surprising? Once again, the answer lies in the fact that subjects' ability to report 100% of the retro-cued row while only being able to report approximately 33% of the entire matrix. Subjects' ability to reproduce 100% of the retro-cued row without consulting any presently available visual information provides evidence that subjects had a conscious representation of 100% of the matrix.

Advocates of overflow also rely on a study by Landman et al. 2003, in which the partial report paradigm exemplified in Sperling (1960) is extended to change blindness stimuli, in order to dispute interpretations of change blindness paradigms that posit a necessary link between attention and

consciousness (Block 2007, Block 2011). In this experiment subjects are briefly shown a circle of eight rectangles in either a horizontal or vertical orientation. Then, after a brief pause, another display of eight rectangles is shown which is identical except for one rectangle that differs in orientation. Ordinarily observers are not able to report this change in orientation. However, when one of the eight rectangles is cued in the interval between the termination of the first display and the onset of the second display, subjects are able to accurately report the change (Landman et. al. 2003). This suggests that subjects possess a detailed conscious representation of the entire first display despite not directing attention to any particular area of it. Importantly, both Sperling (1960) and Landman et al. (2003) suggest that not only are subjects conscious that there *are* letters or rectangles, but that they are conscious of the identity or orientation of those letters or rectangles, meaning that their conscious representations are in fact reasonably rich in detail.

Another experiment that supposedly shows a dissociation between phenomenology and attention uses an object-substitution masking paradigm (Woodman & Luck 2003). In this experimental paradigm the presence of a stimulus is masked by quickly replacing it with another object, leaving subjects unable to report the presence of the first stimulus. In this particular experiment, the display consists of a target stimulus (e.g. a square) surrounded by four small dots and a large number of distracting stimuli (e.g. many triangles) (Woodman & Luck 2003). All of the stimuli disappear after a brief exposure and subjects are asked to report whether or not the target stimulus is present in the display. The experimenters argue that lateralized measures of brain activity in the visual cortex while viewing the display show that subjects did in fact perceive the target stimulus.[5] And, as predicted by the brain activity, when the stimuli all disappear at the same time subjects are relatively accurate in reporting the presence of the target stimuli.

However, in delayed offset trials in which the four small dots surrounding the target stimulus persist after the rest of the stimuli disappear, accuracy in determining whether or not the target stimulus was present is impaired relative to trials in which the target and the dots disappear at the same time, while the lateralized visual cortex activity that is taken to indicate phenomenal consciousness remains the same. Importantly, in the delayed offset trials, by the time attention reaches the target it has already disappeared with only the four dots remaining. According to the experimenters' interpretation, attending to these four dots overwrites whatever phenomenal representation subjects had of the target stimulus (Woodman & Luck 2003). The delayed offset trials are an example of what is plausibly interpreted as visual consciousness as measured by brain activity in the absence of attention as measured by verbal report. Put differently, although subjects were unable to attend to and report the presence of the target, lateralized brain activity shows that it must have been phenomenally conscious.

Why Posit Phenomenal Consciousness?

The three preceding arguments are taken to support the existence of representations that are phenomenally conscious in the absence of attention. In this section I will argue that phenomenal consciousness is not needed in order to formulate viable interpretations of the empirical data. That is, representations that are phenomenally unconscious until they are attended to can perform the same explanatory function as phenomenally conscious representations for each of the aforementioned experimental paradigms.

For example, positing consciousness beyond attention is not the only way to interpret the Sperling partial report paradigm. In particular, there is little reason to think that subjects' representation of the matrix before the retro-cue is phenomenally conscious. Perhaps subjects are able to report the cued rows because they have an unconscious representation of the identities of the letters. The retro cue then causes subjects to attend to the relevant portion of that representation, pulling the identities of those letters into (phenomenal and access) consciousness (Cohen & Dennett 2011). The advocate of overflow might respond with the fact that subjects reported that they experienced more than they could report. Thus, subjects were phenomenally conscious of the whole matrix despite only directing attention to the retro-cued row. However, this reply does not account for the fact that in order to report their experience of the whole matrix, subjects must have had some access to this experience. Perhaps they only saw that there were objects that *could potentially* be identified *if* attention were directed at them, which does not implicate overflow. I will develop this reply further in the next section. For now, all I mean to establish is that there is a viable way to interpret the Sperling paradigm without positing overflow. Thus, the Sperling experiment is not decisive evidence in favor of overflow.

Neither does the change blindness paradigm in Landman et al. (2003) decisively support overflow in the way Block and others think it does (Block 2007). The cue that appeared in between the change might not only have cued subjects to which rectangle in the second display to focus on, but also have acted as a retro-cue, causing participants to attend to an area of their representation of the previous display which is stored in iconic memory. As with the Sperling paradigm, there is no reason to suppose that their representation of the first display is conscious at the point of the cue. The results are just as easily explained by the cue pulling the representation of the rectangle into phenomenal consciousness in virtue of attention being directed at it.

The object-substitution masking paradigm is also susceptible to alternative, non-overflow interpretations. The experimenters establish that attention is directed to the area of the target only after the target disappears (Woodman & Luck 2003). Thus, the delayed offset trials might result in lower accuracy because attention only reaches the area of the target stimulus after it has already disappeared. Since the dots have a delayed offset, subjects attend to the dots instead. Attending to the

dots allows the subject to consciously perceive them while overwriting any representation of the target stimulus due to the working memory limit. What reason do we have to believe that this experiment is decisive in showing that the target stimulus was *consciously* perceived? After all, conditions such as blindsight show that visual information processing can take place without accompanying phenomenal consciousness.[6] One might object that the brain-imaging results of activity in the visual cortex implies that subjects were phenomenally conscious of the target. However, the fact that there was brain activity correlated with perception is still ambiguous between conscious and unconscious perception. None of the reviewed experimental paradigms are able to tease apart these two possible interpretations. At best, overflow is merely one way to explain the data.

Ironically, Block and other proponents of overflow have fallen prey to the very confusion between phenomenology and function that the distinction between phenomenal and access consciousness was introduced to correct. Proponents of overflow seem to implicitly endorse the bridge principle that phenomenal consciousness is present wherever we are able to perform well on visual tasks. For example, on their view, subjects' ability to report the retro-cued rows in the Sperling experiment is a good indicator of the presence of phenomenal consciousness. But, of course, a philosophical zombie without *any* phenomenology would have this same ability. So, this bridge principle is implausible.

The ability of phenomenally unconscious representations to explain the partial report phenomena finds support in the distinction by Dehaene et al. (2006) between preconscious and (phenomenally) conscious visual processing. Preconscious visual processing occurs when visual information that is unaccessed and processed below the level of phenomenal consciousness can still be accessible even after the stimulus disappears. Thus, it is possible for visual information about the Sperling matrix to be processed outside of attention and without phenomenal consciousness and for the retro-cue to then cause that unconscious representation to be accessed and pulled into working memory. This interpretation finds further support in a study that retro-cued attention after a stimulus has already disappeared (Sergent et al. 2013). Like the paradigms mentioned above, the retro-cue enhances subjects' performance in identifying the target stimulus. However, subjective ratings of stimulus visibility support an interpretation in which attention that is directed post-stimulus actually generates a phenomenally conscious perception of the stimuli, rather than merely improving memory of an already phenomenally conscious percept (Sergent et al. 2013).

At this point I should clarify that what I have said so far is not meant to disprove the possibility of consciousness beyond attention. The partial-report, change blindness, and object-substitution masking experimental paradigms might be explained as instances of overflow. However, their experimental design is, unfortunately, ambiguous on this matter. It is uncontroversial that in order for subjects to be as successful as they were, information processing must have taken place beyond focal attention.[7] However, I have shown that it is not clear whether this information processing resulted in

phenomenology or not. This is crucial: the fact that phenomenal consciousness cannot necessarily be identified with function (i.e., the ability to perform well in the experiments described above) undercuts any motivation for accepting overflow. Indeed, there are studies that lend *prima facie* support to the competing, non-overflow interpretation. In the next section, I will answer a potential objection to my argument here while laying out a positive, yet tentative, proposal for the relationship of attention and consciousness that can account for the empirical data and the intuitions philosophers like Block rely on in their arguments for overflow without positing consciousness beyond attention.

Seeming Richness and the Feature-Integration Theory of Attention

So far I have argued that unconscious representations have the same explanatory power as conscious representations in interpreting experiments commonly cited to support consciousness without attention such as Sperling (1960), Landman et al. (2003), and Woodman & Luck (2003). However, the proponent of overflow might object that I have ignored one of the most important pieces of evidence: introspective reports of detailed and complete phenomenology. Remember that subjects in the Sperling paradigm "commonly assert that they can *see* more than they can *report*" (Sperling 1960, emphasis original). In the Landman et al. (2003) paradigm it seems absurd to assert that subjects literally did not see anything in the first display until the retro-cue. Surely all the subjects consciously experienced an array of rectangles, even if they could not report each of their particular orientations. This intuition of rich phenomenology probably lies behind experimental findings that subjects systematically overestimate their performance in change blindness paradigms (Levin et. al. 2000). This is not to mention that in everyday experiences it seems as though we are phenomenally conscious of much more than we could ever hope to report. It certainly seems as if I am not blind to the area outside of my focal attention. Perhaps when taken together, introspective reports of phenomenal richness, coupled with subjects' performance in the experiments discussed above, are evidence enough for consciousness without attention.

Yet, Block and others who rely on the intuition of seeing more than one can report fail to recognize the implications of the fact that subjects in the Sperling experiment were able to report that they saw *something*. The mere fact that subjects were able to report that they had a rich perceptual experience of the entire display shows that the matrix was not entirely inaccessible to them. This is because the ability to introspectively report entails access and therefore can never be used as evidence for inaccessible phenomenal states. By relying on access to determine subjects' phenomenology, Block has failed to show a dissociation between these two types of consciousness. For this stumbling block to be avoided a phenomenal representation must be discovered that is unable to be accessed via introspective report. It is for this very reason that the complicated experimental paradigms described above were crafted in the first place.

Taking subjects' reports at face value, it is important to tease apart precisely what sort of conscious phenomenology the subjects in the Sperling experiment had of the array. David Papineau distinguishes between scene and item phenomenology (Papineau 2007). Papineau calls phenomenology of the display as a whole without detailed representations of the individual letters scene phenomenology. The subjects' reports of seeing the whole display indicate that they had scene phenomenology. What is at issue is whether or not subjects had rich phenomenology of each of the letters in the display. Papineau calls this item phenomenology (Papineau 2007). Not being able to report an experience does not mean it is inaccessible (presumably language deficits do not affect consciousness). However, being able to report something does entail access. Subjects clearly had access to their experience of the array as a whole, given their reports. Their inability to report the identities of every letter in the array, on the other hand, shows that they lack access to the specific alphanumerical identities of the symbols. So, subjects had scene phenomenology and scene access without item access. Whether phenomenal overflow occurs hinges on whether subjects had item phenomenology.

The distinction between scene and item phenomenology is quite intuitive. For example, you can probably recall a case where you laid your eyes upon a scene briefly and saw various colors, shapes, lighting, motion, and depth. Only afterwards were you able to piece these together into distinct objects. You might also imagine viewing an impressionist painting and at first only being able to see the various brushstrokes (the scene) before being able to understand how they form the objects (the items) that the artist intended. The converse can also be true; one might have access to high-level representations without access to low-level representations. An example of this might be getting the gist, or meaning, of a picture without seeing the individual perceptual contents that made it up. That these examples are hard to come by suggests just how closely phenomenology and access map on to each other in most situations.

So scene and item phenomenology can be conceptually distinguished, but can they truly come apart in practice? If they can, what cognitive mechanisms lie behind this process? Empirical results largely support dissociations between different levels of phenomenology (Kouider et al. 2007). On a plausible but controversial interpretation, patients with type-2 blindsight have degraded phenomenology such that they are able to report that they saw *something* without being aware of any determinate features of their phenomenology (Brogaard 2015, Overgaard et. al. 2008, for an opposing view see Weiskrantz 2009). This is plausibly interpreted as an instance of scene phenomenology without item phenomenology. More empirical backing for this distinction is found in patients with integrative visual agnosia. Integrative agnosiacs are "impaired at search tasks that require the binding of visual elements in a spatially parallel manner across a field containing multiple stimuli" (Behrmann & Kimchi 2003). That is, although they may be aware of the various features present in a visual scene they are unable to integrate those features into coherent objects. When patients with integrative

agnosia are shown two differently colored letters and are asked to name the color of the first letter they see, they will often report the color of the other letter (Treisman 1998). This is taken to show that agnosiac patients are conscious of the various colors and letters that are present, but are impaired at binding the colors and letters together into objects.

Importantly, these very same binding errors occur in neurotypical subjects who are prevented from focusing attention on any of the colored letters in particular through the use of brief presentation and distractor objects (Treisman 1998). The importance of attention for binding features is also suggested by the asymmetry in search times when looking for a shape with an added feature (searching for a Q among O's) compared to a shape that lacks that same feature (searching for an O among Q's). The latter requires binding the tails to each Q while the former merely requires searching for the one feature that stands out (Treisman 1998). It is beyond the scope of this paper to review the extensive evidence for the feature-integration theory of attention. However, the evidence reviewed above suggests that it is a live, and likely, hypothesis that attention is necessary for binding features into objects.

The feature-integration theory of attention empirically fleshes out the conceptual distinction between scene and item phenomenology, especially in the context of the Sperling (1960) and Landman et al. (2003) experiments. In particular, we can operationalize the pre-attentive *features* as scene phenomenology and perception of the post-attentive *objects* as item phenomenology.[8] The feature-integration theory of attention can help to make sense of subjects' performance on experiments such as Sperling (1960) along with their intuitions of rich phenomenology without positing consciousness beyond attention. During the brief display of the Sperling matrix subjects are able to make out the various features present in the matrix. This explains subjects' ability to characterize the display as containing only letters, as well as their intuition of seeing more than they can report. Once the retro-cue is administered attention is directed to an area of the matrix and the features within that area are bound into objects before the representation of the matrix decays. This explains subjects' abilities to accurately report any cued row without consulting a rich, phenomenally conscious representation of the entire matrix. The representation only becomes phenomenally conscious and richly detailed once the retro-cue directs attention to a portion of it.

This interpretation finds support from an experiment which replicated the Sperling paradigm while also introducing "pseudo-letters" in non-cued rows (de Gardelle et al. 2009). Subjects were led to expect either real letters or very different symbols (such as a star or a face) in the array. Using free subjective report researchers found that not only was it the case that "pseudo-letters were hardly noticed" but often "participants actually tended to perceive the pseudo-letters as real letters," or, at least, subjects *reported* seeing pseudo-letters as real letters (de Gardelle et al. 2009). Why is this significant? Imagine for example, that while reading an ordinary text you saw this symbol: "Ƨ", instead of "S" in the middle of a word. This pseudo-letter would undoubtedly stand out to you as being out

of the ordinary. We should expect the same of the de Gardelle et al. (2009) subjects if they had rich phenomenology. That subjects were unable to do this indicates that their phenomenology cannot have been as richly detailed as Block posits.

Importantly for my account, the feature-integration theory of attention can explain why pseudo-letters were often mistaken for their actual counterparts. Pseudo-letters share nearly all of the same low-level features as actual letters. For example, Ƨ is made up of the same curved lines as S. Since pseudo-letters only occurred outside of cued rows, subjects never bound the features together. Then, because subjects were primed to expect only real letters, they inferred that the presence of the features indicated the presence of a real-letter instead of a pseudo-letter. Thus, de Gardelle et al. (2009) demonstrates an illusory conjunction outside of attention in a very similar way to Treisman (1998). This suggests that subjects did not so much directly perceive the Sperling matrix as reconstruct it after the cue based on the guiding assumption that the features present should combine to form letters. Sperling (1960) is not a case of perception without attention but rather a case of attention (coupled with top-down assumptions) *guiding* perception.[9]

The defender of overflow might object that I have conceded their point. Since my account entails that conscious perception of *features* can occur outside of focal attention, my argument has in fact established the possibility of scene phenomenology overflowing attention. This is a weaker form of overflow than the kind of rich, item phenomenology without attention that most proponents of overflow, including Block, argue for, but it is a potential counterexample to my argument nonetheless. In what follows, I will give three reasons for why this supposed case of overflow is not a counterexample to my claim that attention is necessary for consciousness.

First, as discussed above, the mere fact that subjects are able to report the fact that they have scene phenomenology beyond merely the cued items entails that they have some sort of access to it, attentional or otherwise. If subjects truly had phenomenology of the display that was not access conscious then they should not be able to report it at all because verbal report indicates the ability to voluntarily and rationally make use of information: the hallmark of accessibility (Block 1995). The fact that they were able to report the mere impression of having seen the whole display shows that their scene phenomenology was accessed.

Second, all the feature-integration theory of attention posits is perception of features outside of *focal* attention. There is converging evidence for the presence of distributed attention throughout most of the visual field (Srinivasan et al. 2009). Distributed attention is weaker than focal attention and picks up different kinds of visual information than focal attention, such as global statistical properties of a scene (Srinivasan et al. 2009). So, the phenomenal awareness of the scene that subjects reported in all of the experimental paradigms under discussion may result from distributed attention to items outside of focal attention (Srinivasan et al. 2009, Brigard & Prinz 2010, Cohen & Dennett 2011). Lending support to this interpretation, Cohen et al. (2011) found that when a distractor task

such as multiple object tracking was sufficiently demanding, inattentional blindness to the scene occurred. The experimenters used natural scenes such as animals or mountains in order to maximize ecological validity and found that scene perception decreased as the demands of the task increased (Cohen et al. 2011). The experimenters interpret this result as showing that studies which claim to show phenomenal consciousness beyond attention have not sufficiently engaged *all* attentional resources. The leftover attentional resources can then attend to the scene in the form of distributed attention (Cohen et al. 2011). It is distributed attention that allows the scene to be perceived along with focal attention which binds features together and allows for the report of specific items in paradigms such as Sperling (1960) and Landman et al. (2003).

Finally, there are a wealth of studies in the vision science literature that suggest that humans can perceive statistical regularities in our environments with very little attentional demand (Cohen et al. 2016, Greene & Oliva 2009, Srinivasan et al. 2009). In other words, just because attention is a necessary condition for consciousness, and just because attention has a limited capacity, does not necessarily mean that our conscious experience is sparse. This doesn't just accord with the intuitions of the Sperling subjects who insisted that they consciously perceived the entire matrix of letters, but it also validates our everyday experience. The idea is that, when presented with a densely speckled hen, instead of consciously representing each speckle individually, we simply represent that there is an *ensemble* of speckles, thereby condensing what would have been many bits of information that would have placed a high demand on our attentional resources into a single bit of information that is not as attentionally demanding. This ability allows us to get the "gist" of a visual scene without representing individual objects and seems to be importantly related and, indeed, may be integral to the formation of scene phenomenology. Focal attention might give us a "high-resolution" picture of the individual speckles, but it does not need to be present in order to represent the ensemble property.

Importantly, phenomenal consciousness of global statistics and ensemble properties does not exceed the limits of attention. As described above, inattentional blindness to the scene occurs as attentional demand increases (Cohen et al. 2011). Additionally, "observers are more accurate at processing multiple ensembles when they are presented sequentially, rather than simultaneously," suggesting that they use up a limited pool of attentional resources (Cohen et al. 2016). Our ability to quickly represent ensembles even with only limited attentional allocation, along with the presence of distributed attention, is consistent with the feature-integration theory of attention and fully explains the intuition of seeing more letters than one can report without positing consciousness beyond attention. So, neither scene nor item phenomenology overflow attention. Each type of phenomenology has distinct attentional mechanisms, focal and distributed; high-resolution and gist, that allow visual scenes to be consciously perceived.

The feature-integration theory elaborated above, coupled with phenomena like distributed attention and perception of ensemble properties can help defuse some of the more recent studies

thought to support overflow. For example, one study replicated the Sperling experiment with colored letters (Bronfman et al. 2014). The level of color diversity was varied both in the cued and the non-cued rows. The experimenters found that subjects could accurately estimate the level of color diversity outside of the cued row with no cost to letter report. This is taken to show that subjects had color phenomenology outside of focal attention and without any addition to working memory load (Bronfman et al. 2014). However, my account can circumvent this interpretation, and indeed, comfortably accommodate these results. This is because high color diversity implicates the presence of a variety of different features. In fact, accurate color diversity judgments could be made based on an awareness of color features alone, which according to the feature-integration theory can be registered pre-attentively. On top of this, color diversity is an ensemble property that can be registered with very little attention necessary (although not without *any* attention). In other words, subjects do not need a rich representation of each individual colored letter in order to accurately estimate color diversity. This study, far from being evidence for overflow, actually demonstrates one way in which we can access a wealth of visual information without recourse to a rich phenomenal representation beyond the reach of attention.

Conclusion

In the preceding sections I have argued that attention is a necessary condition of consciousness. The various arguments that have been put forward for phenomenal overflow are inconclusive insofar as they can be reinterpreted without positing overflow. The feature-integration theory of attention, coupled with distributed attention and perception of ensemble properties, can account for the intuition of phenomenal richness as well as subjects' performance in paradigms such as Sperling (1960) and Landman et al. (2003) without positing consciousness beyond attention. Thus, arguments for phenomenal overflow are at best inconclusive or at worst ignore important experimental findings concerning attention and therefore do not stand up to empirical scrutiny. There are important questions that future research must grapple with such as whether phenomenal overflow poses a methodological threat to a science of consciousness (Block 2011, Cohen & Dennett 2011). This paper has focused on synthesizing and interpreting the vast empirical literature on attention and consciousness and has tentatively proposed that attention as a necessary condition for consciousness best accommodates the evidence currently available.

Notes

[1] There is an interesting literature on whether attention is *sufficient* for consciousness (Brigard & Prinz 2010, Kentridge et al. 1999, Kentridge et al. 2008, Mole 2014). I will limit my discussion in this paper to the question of necessity.

[2] Indeed, as I will go on to show, Block's arguments for overflow hinge on particular details about certain experiments, all of which are in the visual modality and employ very brief exposures, thus calling into question their generalizability.

[3] One famous experiment found that the working memory limit is, in fact, approximately seven (Miller 1956). One potential reason for this discrepancy between the Miller (1956) and the Sperling (1960) working memory findings is the relatively much briefer exposure to the Sperling display.

[4] A crucial part of understanding the Sperling experiment that gets overlooked in the literature on overflow is that language (including the representation of letters) is processed in specialized language modules. In other words, letter cognition and recognition is not a purely perceptual task, it relies on other cognitive mechanisms. Therefore, using letters in the Sperling task may not be the best way to prove phenomenal overflow, which is supposed to be a purely perceptual phenomenon. There is not enough research in this domain to sketch out the implications of this for overflow.

[5] By lateralized brain activity I mean differential activation in the visual cortex that is contralateral to the side of the visual field that the target stimulus was present on. This differential visual cortex activation provides an implicit measure of visual consciousness of the target.

[6] Ned Block's argument that blindsight constitutes a deficit of both phenomenal *and* access consciousness is well taken (Block 1995). However, all I need to show is that visual information can be processed (not accessed) in the absence of phenomenal consciousness in order to draw a dissociation between visual information processing and conscious perception.

[7] More on the distinction between focal and distributed attention later.

[8] From here on I shall use scene phenomenology to mean perception of unbound features and item phenomenology to mean perception of bound objects.

[9] Some evidence (Blackmore et al. 1995) suggests that with each attentional refocus that accompanies saccadic eye movements we constantly have to reconstruct items out of an ever-changing scene, with very limited transsaccadic memory.

Bibliography

Behrmann, M., and R. Kimchi. "What Does Visual Agnosia Tell Us about Perceptual Organization and Its Relationship to Object Perception?" *Journal of Experimental Psychology: Human Perception & Performance* 29.1 (2003): 19-42.

Blackmore, S., G. Brelstaff, K. Nelson, and T. Troscianko, T. "Is the Richness of Our Visual World an Illusion? Transsaccadic Memory for Complex Scenes." *Perception* 24.9 (1995): 1075-1081.

Block, N. "On a Confusion about a Function of Consciousness." *Behavioral and Brain Sciences* 18.2 (1995): 227.

Block, N. "Consciousness, Accessibility, and the Mesh between Psychology and Neuroscience." *Behavioral and Brain Sciences* 30 (2007): 481-548.

Block, N. "Perceptual Consciousness Overflows Cognitive Access." *Trends in Cognitive Sciences* 15.12 (2011): 567-575.

Brigard, F., and J. Prinz. "Attention and Consciousness." *WIREs: Cognitive Science* 1.1 (2010): 51-59.

Brogaard, B. "Type 2 Blindsight and the Nature of Visual Experience." *Consciousness and Cognition* 32 (2015): 92-103.

Bronfman, Z. Z., N. Brezis, H. Jacobson, and M. Usher. "We See More Than We Can Report: 'Cost Free' Color Phenomenality Outside Focal Attention." *Psychological Science* 25.7 (2014): 1394-1403.

Carruthers, P. "Block's Overflow Argument." *Pacific Philosophical Quarterly* 97(2015).

Cohen, M., G. Alvarez, and K. Nakayama. "Natural-Scene Perception Requires Attention." *Psychological Science* 22.9 (2011): 1165-1172.

Cohen, M., and D. Dennett. "Consciousness Cannot Be Separated from Function." *Trends in Cognitive Sciences* 15.8 (2011): 358-364.

Cohen, M. A., D. Dennett, and N. Kanwisher. "What Is the Bandwidth of Perceptual Experience?" *Trends in Cognitive Sciences* 20 (2016): 324-335.

De Gardelle, V., J. Sackur, and S. Kouider. "Perceptual Illusions in Brief Visual Presentations." *Consciousness and Cognition* 18 (2009): 569-577.

Dehaene, S., J. Changeux, L. Naccache, J. Sackur, and C. Sergent. "Conscious, Preconscious, and Subliminal Processing: A Testable Taxonomy." *Trends in Cognitive Sciences* 10.5 (2006): 204-211.

Dretske, F. "What Change Blindness Teaches about Consciousness." *Philosophical Perspectives* 21 (2007): 215-230.

Greene, M., and A. Oliva. "Recognition of Natural Scenes from Global Properties: Seeing the Forest without Representing the Trees." *Cognitive Psychology* 58.2 (2009): 137-176.

Kentridge, R. W., C. A. Heywood, and L. Weiskrantz. "Attention without Awareness in Blindsight." *Proceedings of the Royal Society B: Biological Sciences* 266.1430 (1999): 1805-1811.

Kentridge, R., T. Nijboer, and C. Heywood. "Attended but Unseen: Visual Attention Is Not Sufficient for Visual Awareness." *Neuropsychologia* 46.3 (2008): 864-869.

Kouider, S., V. Gardelle, J. Sackur, and E. Dupoux. "How Rich Is Consciousness? The Partial Awareness Hypothesis." *Trends in Cognitive Sciences* 14.7 (2010): 301-307.

Landman, R., H. Spekreijse, and V. Lamme. "Large Capacity Storage of Integrated Objects before Change Blindness." *Vision Research* 43 (2003): 149-164.

Levin, D., N. Momen, S. Drivdahl, and D. Simons. "Change Blindness Blindness: The Metacognitive Error of Overestimating Change-detection Ability." *Visual Cognition* 7 (2000): 397-412.

Miller, G. A. "The Magical Number Seven, Plus or Minus Two: Some Limits on Our Capacity for Processing Information." *Psychological Review* 63.2 (1956): 81-97.

Noë, A., L. Pessoa, and E. Thompson. "Beyond the Grand Illusion: What Change Blindness Really Teaches Us about Vision." *Visual Cognition* 7 (2000): 93-106.

Overgaard, M., K. Fehl, K. Mouridsen, B. Bergholt, and A. Cleeremans. "Seeing without Seeing? Degraded Conscious Vision in a Blindsight Patient." *PLoS ONE* 3.8 (2008).

Papineau, D. "Reuniting (Scene) Phenomenology with (Scene) Access." *Behavioral andBrain Sciences* 30 (2007).

Schwitzgebel, E. *Perplexities of Consciousness*. Cambridge, Mass: MIT Press, 2011.

Sergent, C., V. Wyart, M. Babo-Rebelo, L. Cohen, L. Naccache, and C. Tallon-Baudry. "Cueing Attention after the Stimulus Is Gone Can Retrospectively Trigger Conscious Perception." *Current Biology* 23.2 (2013): 150-155.

Siegel, S., and N. Silins. "Consciousness, Attention, and Justification." In *Scepticism and Perceptual Justification*. Oxford: Oxford University Press, 2014.

Simons, D., and D. Levin. "Change Blindness." *Trends in Cognitive Sciences* 1.7 (1997): 261-267.

Sligte, I., H. Scholte, and V. Lamme. "Are There Multiple Visual Short-Term Memory Stores?" *PLoS ONE* 3.2 (2008).

Sperling, G. "The Information Available in Brief Visual Presentations." *Psychological Monographs: General and Applied* 74 (1960): 1-29.

Srinivasan, N., P. Srivastava, M. Lohani, and S. Baijal. "Focused and Distributed Attention." *Progress in Brain Research: Attention* 176 (2009): 87-100.

Treisman, A. "Feature Binding, Attention and Object Perception." *Philosophical Transactions of the Royal Society B: Biological Sciences* 353 (1998): 1295-1306.

Weiskrantz, L. "Is Blindsight Just Degraded Normal Vision?" *Experimental Brain Research* 192.3 (2009): 413-416.

Woodman, G. and S. Luck. "Dissociations among Attention, Perception, and Awareness during Object-substitution Masking." *Psychological Science*. 14.6 (2003): 605-611.