Getting Clear on the Challenge[1]

Commentary on Gross and Flombaum "Does Perceptual Consciousness Overflow Cognitive Access? The Challenge from Probabilistic, Hierarchical Processes"

Nico Orlandi and Aaron Franklin – UC Santa Cruz

Gross and Flombaum (G&F) offer a reinterpretation of three studies that are standardly taken to demonstrate both the existence of different types of memory stores of declining capacity, and a distinction between phenomenal consciousness and access consciousness (Block 1995, 2007, 2011). We are unable to dwell on the many interesting aspects of Gross and Flombaum's paper. Our comments focus on G&F's positive proposal, leaving aside whether they are successful in raising problems for competing explanations.

Our main worry is that it is not clear what aspects of the view that Gross and Flombaum propose do explanatory work, and derivatively how their ideas explain the results. The authors present a complex position. They favor a certain way of understanding perceptual processing, as well as a model of visual memory. These two aspects of their view share common elements – for example, they have in common the notion of probabilistic representation – but it is often hard to follow what aspect is central to explanation. We divide the commentary into three sections corresponding to the three main studies discussed in the original article.

1. Letters

In Sperling's now classic experiment (1960), participants are flashed with an array of letters in three rows and four columns (12 in total), and they are asked to recall what letters they saw. Sperling found that participants so instructed – although remembering having seen many of the letters – could report only about 4 or 5. Sperling then used a partial-report procedure in which participants were cued by an auditory beep to report a randomly selected row. If the beep was played right when the letters disappeared, participants could report nearly all the letters from the cued row. This suggests that participants saw more letters than what they could report in the non-cued condition. If the beep was played after a delay of 500 ms, reports for a given row reduced greatly, resulting in an estimate of about 1.5 letters per row.

This study is standardly taken to reveal the presence of a high-capacity iconic memory (Neisser, 1967) that dissipates within a half a second. This type of memory is different from working memory that has less capacity, and where things are stored after stimulus offset. The study has also been used to argue that phenomenal consciousness overflows access consciousness (Block 2011). Perceivers are sometimes perceptually conscious of things that are not available to report.

In contrast with this standard way of interpreting the results, Gross and Flombaum propose an alternative that does not commit to a difference between phenomenal and access consciousness, and does not commit to the presence of different memory stores of declining capacity. G&F propose the idea – fairly widespread in vision science – that visual perception is an inferential and probabilistic process that constructs probabilistic representations of

[1] Thanks to Jacob Beck for comments on an early draft of this commentary.

distal scenes from noisy transduced signals. Such representations are stored in a single-resource memory store. Performance in response to a specific query involves "sampling" from the probabilistic representations stored in visual memory (p. 6). Sampling consists in selecting a specific hypothesis from a space of representations with a certain probability (p. 23).

Probabilistic representations, in this context, are "assignments of credences over a space of possibilities concerning the distal scene" (p. 6). Using an example from later in the article, probabilistic representations have the form: <*F, there*$_i$*, .7*>, <*T, there*$_i$*, .1*>, with the number indicating the subjective probability of the letter "F" or of the letter "T" being present. A non-probabilistic representation would instead look like this: <*F, there*$_i$> (p. 30).[2]

How does this picture explain the Sperling results? In some passages, Gross and Flombaum stress the idea that participants in visual memory experiments "can report only 3 to 5 objects correctly because that's all their visual system can *confidently* recover, on average." (p. 17). This indicates that the explanation centers on the notion of probabilistic representation. The visual system, for lack of time, fails to assign high probability to the specific letter representations it has processed. In this view, subjects only report letters that are represented as probable, or as passing a certain probability threshold, ignoring the representations that are improbable. When a cue is introduced, the cue focuses resources on a given row, helping the representations of the letters in that row to acquire high probability.

In other passages, Gross and Flombaum make use of the idea that the visual system does not have sufficient time, before the cue, to form "highly detailed" (p.14) and "rich, specific" representations (p. 15). In an inferential framework – they say – representations of letters (high-level features) are inferred from representations of low-level features – for example, oriented edges (p. 7) – so recovering representations of *any letters at all* takes time and resources (p. 19). The role of the cue is not to transfer representations one already has from one store to another (p. 14), but to help the system infer the representations it needs for reporting (pp. 25-26).

So, G&F present two different ways of explaining the Sperling results. In the first explanation, the visual system represents the letters in the un-cued display, but for many of them, it assigns a low probability – too low to make it to report. In the second explanation, the idea is that the visual system, for lack of time, fails to recover representations of specific letters *at all*. For the majority of the letters in the un-cued display, the visual system only represents the letters' low-level features – <*oriented edge*, *there*$_i$>.[3]

It is not clear which of these two ways of accounting for the studies is Gross and Flombaum's way. Moreover, both ways have problems. Consider first the latter strategy, the one that centers on the inferential nature of perception. For lack of time, the visual system fails to infer letter representations at all. We have two worries about this view. One is that the inferential claim, while seemingly important, is superfluous to the explanation. If vision acted in a purely associative manner – associating low-level sensory states to create percepts

---

[2] There are actually two ways of understanding the notion of probabilistic representation. One where the probabilities are part of the content of the representation, the other where the probabilities attach to the attitudes that subjects have towards given contents (Munton 2016). Following Gross and Flombaum's own notation, we assume that the probabilities are part of the content of perceptual representations.

[3] There is a further possibility where the low-level features are represented probabilistically. We leave this option aside to simplify the discussion.

– its processing would take time. So long as the process takes time, and is enhanced by attention, we have an explanation of the results that is silent on the nature of visual processing.

The second worry is phenomenological. Gross and Flombaum admit that their view has the burden of explaining why subjects report seeing more letters than what they can remember (p 11). Block does not have this problem since he thinks that iconic memory contains representations of (almost) all of the letters in the display. G&F, however, cannot appeal to unreported letter representations in this interpretation of their view. This is because participants have not seen (or represented) unreported letters. They have only represented letter fragments. So both their seeming to see the letters, and their remembering having seen them are illusory.

This interpretation of Gross and Flombaum's position would, in fact, predict that subjects should report *making a determination about what they have seen*, not *remembering having seen more letters than what they can report*. If a participant sees an *F* and then tries to remember what she saw, she is doing something different than trying to determine, in memory, what letter she saw from fragments of the letter. The subjects in Sperling's experiment are doing the former, while Gross and Flombaum's model suggests that they should be doing the latter.

Now, there are reasons to think that G&F do not rely on the inferential nature of vision to explain the results. They concede that not all of the points they make in interpreting the studies require all aspects of their proposal (p. 6), and they seem open to giving up the hierarchical aspect of their view – that is, the claim that vision uses low-level feature representations. Further, when discussing memory, they say that the contents of visual memory are not individual letters or other feature values, but probability distributions (or densities) over *those entities* (p. 22). According to G&F, the capacity limits in memory concern the *precision* of representations, not the *number* of items that can be represented. Concerning number, "no limitation may be assumed at all!" (p. 31).

So let's consider, instead, the explanation that makes reference to probabilistic representations of letters. In the Sperling studies, the visual system represents the letters in the display, but for many of them, it assigns a low probability – too low to make it to report. We again have two worries. One is that, in this view, the visual system represents a whole lot of letters, many more than the letters that are reported. This seems to be just Block's position – the main difference being that iconic memory stores probabilistic representations instead of discrete ones. [4]

The other worry is that this type of explanation is subject to the same criticism that Gross and Flombaum wage against competing views (p. 15). According to them, performance limitations and time constraints make it implausible that visual systems would form "highly detailed" (p.14) and "rich, specific" representations of uncued letters that are then stored in memory either consciously (Block 2011) or unconsciously (Phillips 2011). But if it is troublesome to form discrete representations of letters, why is it any less troublesome to form probabilistic representations of letters? Why is representing *F,* any harder than representing *probably F* (or *F to a probability of .7*)? Considerations of performance limitations and time constraints should lead one to conclude that the visual system does not go to the trouble of representing letters (probabilistically or otherwise), but only fragments of letters.

In sum, more clarity is needed to assess G&F's challenge in the Sperling studies.

[4] See also the commentary by Jacob Beck for a similar point.

2. Orientation

We now turn to experiments that require appeal to another aspect of Gross and Flombaum's proposal, the view that visual memory consists in a store of probabilistic representations of the form $<F, \textit{there}_i, .7>$, and that performance involves "sampling" from this store.

In the experiments by Lamme and colleagues, subjects are initially presented with one display containing various segments oriented in different ways (Vandenbroucke et al. 2011). In a second display, some of the segments differ in orientation. Subjects have to signal whether there is a change between the two displays. Like in the Sperling studies, performance is modulated by cues. Subjects perform significantly better in the early cue condition, where a cue as to where the change occurs, is presented 10ms after the first display. They perform worse in the retro cue condition, where the cue comes between the two displays (for example 1000ms after the first and 500ms before the second). They perform even worse in the post cue condition, when the cue comes on along with the second display. Because of this difference in performance, the studies have been thought to indicate the existence of a third memory store – fragile memory – intermediate between iconic and working memory.

On their analysis of this study, Gross and Flombaum focus mostly on rejecting the memory-store interpretation, but we think more clarity is needed in explaining how G&F's own proposal explains the results. Good performance in the early cue condition is presumably explained by reference to the cue reducing the scope of relevant inputs – and thus the computational complexity (p. 33). Particularly puzzling, however, is the difference in performance between the retro and the post cue conditions. The retro cue appears too late to affect perceptual processing. From other studies, we know that it takes less than 400ms and, generally, only 200ms to obtain perceptual completion (Sekuler and Palmer 1992). It also takes around 500ms to experience perceptual switching – for example, seeing the Necker cube in two different ways (Orlandi 2012). When the retro cue comes 1000ms after the first display, this is beyond the window of perceptual processing. If this were a matter of perception only, then, in this retro condition, subjects should perform comparably to subjects in the Sperling studies whose cue comes after 500ms of stimulus off-set (reporting only 1.5 letters per row). But this is not the result.

Gross and Flombaum have to appeal to memory in this case. Memory, according to them, contains probabilistic representations that are sampled when subjects are cued. This story, however, does not explain the discrepancy in performance. The cue may guide focal attention to the relevant positions in memory (p. 25), but why would the timing of the cue make a difference if working memory is a single store? Gross and Flombaum say: "(…) by re-deploying attentional resources, the retro-cue can selectively protect corresponding representations of items in the first display from degradation and interference. The post-cue comes too late to have this effect." (p. 33) We are not sure why the post-cue would be too late, while the retro cue is not. Both seem to be too late when it comes to perception, and to be roughly equivalent when we consider memory as a single store.

3. Color

In the final study discussed in this article, participants are briefly presented with an array of letters (Bronfman et al. 2014). Unlike in the Sperling study, the letters are *variously colored*, and participants are asked to report only one of the letters in the cued row. In addition, participants are asked to report on color diversity. On some trials, the letters on any given row are highly different in color. In others, their color diversity is low, ranging between 6 adjacent colors in the color wheel. In one condition, subjects are asked to report on color diversity of the cued row, while in a second condition, the report is on the un-cued rows. The central result is that subjects are typically able to report the letter correctly, *but also* to report accurately the color diversity of both cued and un-cued rows. This experiment is taken to further support the thesis that phenomenal consciousness overflows access consciousness. Subjects are consciously aware of more than what they can report.

Gross and Flombaum again focus on raising problems for this interpretation. But the studies seem problematic for their view of perceptual processing and of working memory. Why should we get – consciously or unconsciously, and probabilistically or discretely – color representations "for free" in these cases? Color perception is a complicated and highly noisy matter, dependent on factors such as the quantity of light reflected by a stimulus, the illumination, and the transmittance properties of the space between the stimulus and the observer. Inferential accounts of vision are thought to be the best models of color perception (Brainard 2009). But, if this is so, why are participants seemingly forming accurate and speedy representations of un-cued colors in a display? Shouldn't their color representations either not be there, or have a particularly low probability such that they do not make it to report?

G&F remark that the results may be explained by participants seeing just 3 of the colors in un-cued rows. (p. 37). They also point out that the studies are compatible with the participants not having seen specific colors, but only generic (or "gist") colors (p. 40). Even then, however, the processing of 3 specific colors, and/or of generic colors takes time and attentional resources. Following G&F's explanation of the Sperling and Lamme experiments, we should expect subjects to represent confidently the colors of the cued rows, but expect them to have no representations of colors at all, or representations with very low probability, of the colors in the un-cued rows.

Further, Bronfman and collaborators explain that when both cued and un-cued rows have the same color diversity – they are "congruent" – participants' accuracy in estimating the color diversity of the non-cued rows is *not different from* their accuracy in estimating the color diversity of the cued row (Bronfman et al. 2014, pp. 1399-1400). So, it is not just that participants represent colors of the un-cued rows, it is also that they seem to represent those colors to the same extent that they represent colors for the cued rows. This is perplexing. In Gross and Flombaum's explanation a cue should make a difference to the probabilistic representations that are produced in perception, or sampled in memory. G&F's view seems therefore to be at odds with these results.

4. Conclusion

Gross and Flombaum present a model of perception and of visual memory that accords with current research in cognitive science. In our commentary, we raised questions concerning the ability of this model to explain a set of studies that are commonly taken to bear on our understanding of memory and of consciousness. G&F's proposal needs to be better spelled out. We are unclear on what features of their view does explanatory work, and on what aspects of the view make it both distinctive, and better than competitors.

Works Cited

Block, N. 1995: On a confusion about a function of consciousness. Behavioral and Brain Sciences, 18, 227-87.

Block, N. 2007: Overflow, access, and attention. Behavorial and Brain Sciences, 30, 530-48.

Block, N. 2011: Perceptual consciousness overflows cognitive access. Trends in Cognitive Science, 12, 567-75.

Brainard, D. H. 2009. Bayesian approaches to color vision. The visual neurosciences, 4. available online at http://color.psych.upenn.edu/brainard/papers/BayesColorReview.pdf.

Bronfman, Z., Brezis, N., Jacobson, H., and Usher, M. 2014: We see more than we can report: 'cost-free' color phenomenality outside focal attention. Psychological Science, 25, 1394-403.

Munton, J. 2016. Visual Confidences and Direct Perceptual Justification. Philosophical Topics, 44(2), 301-326.

Phillips, I. 2011. Perception and iconic memory: what Sperling doesn't show. Mind & Language, 26, 281-311.

Neisser, U. 1967: Cognitive Psychology. Englewood Cliffs, NJ: Prentice-Hall.

Orlandi, N. 2012. Visual switching: The illusion of instantaneity and visual search. Review of Philosophy and Psychology, 3(4):469{480.

Sekuler, A. B. and Palmer, S. E. 1992. Perception of partly occluded objects: A microgenetic analysis. Journal of Experimental Psychology: General, 121(1):95.

Sperling, G. 1960: The information available in brief visual presentations. Psychological Monographs, 74.

Vandenbroucke, A., Sligte, I., and Lamme, V. 2011. Manipulations of attention dissociate fragile Visual Short-Term Memory from Visual Working Memory. Neuropsychologia, 49, 1559-68.