CellPress

Opinion

What is the Bandwidth of Perceptual Experience?

Michael A. Cohen,[1,*] Daniel C. Dennett,[2] and Nancy Kanwisher[1]

Although our subjective impression is of a richly detailed visual world, numerous empirical results suggest that the amount of visual information observers can perceive and remember at any given moment is limited. How can our subjective impressions be reconciled with these objective observations? Here, we answer this question by arguing that, although we see more than the handful of objects, claimed by prominent models of visual attention and working memory, we still see far less than we think we do. Taken together, we argue that these considerations resolve the apparent conflict between our subjective impressions and empirical data on visual capacity, while also illuminating the nature of the representations underlying perceptual experience.

Perception: Rich or Sparse?

The moment we open our eyes, we experiences a vast, richly detailed visual world extending well into the periphery [1,2]. However, numerous experimental results indicate that the bandwidth of human perception is severely limited. Findings from change blindness and inattentional blindness demonstrate that much of the available visual information goes unnoticed [3]. Direct estimates of the capacity of visual **attention** (see Glossary) and working memory reveal that surprisingly few items can be processed and maintained at once [4,5]. These results raise a natural question: why do we think we see so much when the scientific evidence suggests we see so little?

One answer to this question is that change blindness and inattentional blindness highlight the limits of mechanisms such as attention and working memory, rather than the limits of conscious perception. According to this view, perception 'overflows' and exceeds the capacity of the cognitive mechanisms needed to access that information [6]. In other words, we consciously perceive more than we can attend, remember, report, or base decisions on [7–11]. Under this view, the neural processes associated with visual **awareness** are separate from those associated with attention, working memory, and explicit report. **Recurrent processing** in sensory cortex supports conscious perception [10], whereas the parietal and prefrontal cortices support the cognitive mechanisms involved in accessing those percepts [12]. According to this framework, there is no tension between our subjective impression of the world and objective measures of human capacity limits because both of these are true. We have a rich experience of the world that cannot be fully captured by the capacity-limited cognitive mechanisms beyond the canonical visual system.

However, contrary to this view, many researchers argue that awareness is intrinsically linked to these cognitive functions and information is not consciously perceived until it is accessed by higher-order systems, such as attention, working memory, and decision-making [13–18]. Rather than link conscious perception with recurrent processing in sensory cortex, this view associates awareness with the parietal and prefrontal cortices [14]. However, for those who endorse this view, the problem remains: how can our impression of a rich visual experience be supported by mechanisms that have strict capacity limits? Put another way, it has been claimed

Trends

Numerous empirical results highlight the limits of visual perception, attention, and working memory. However, it intuitively feels as though we have a rich perceptual experience, leading many to claim that conscious perception overflows these limited cognitive mechanisms.

A relatively new field of study (visual ensembles and summary statistics) provides empirical support for the notion that perception is not limited and that observers have access to information across the entire visual world.

Ensemble statistics, and scene processing in general, also appear to be supported by neural structures that are distinct from those supporting object perception. These distinct mechanisms can work partially in parallel, providing observers with a broad perceptual experience.

Moreover, new demonstrations show that perception is not as rich as is intuitively believed. Thus, ensemble statistics appear to capture the entirety of perceptual experience.

[1]Department of Brain and Cognitive Sciences, McGovern Institute for Brain Research, Massachusetts Institute of Technology, Cambridge, MA, USA
[2]Center for Cognitive Studies, Department of Philosophy, Tufts University, Medford, MA, USA

*Correspondence: michaelthecohen@gmail.com (M.A. Cohen).

that 'Introspectively, consciousness seems rich in content. . .From the third-person perspective of the behavioral scientist, however, consciousness is rather miserable' ([10] p. 205).

We argue here that, even though conscious perception is limited by cognitive mechanisms such as attention and working memory [3], it is not 'rather miserable', and the visual information observers have access to is not at all sparse. To make this argument, we discuss a variety of recent results demonstrating that people can encode and remember considerably more than just a few items. First, we examine empirical findings from a relatively new field of study: visual **ensembles and summary statistics** [19]. The key idea here is that the visual system exploits the redundancy found in real-world scenes to represent a large amount of information, often extending into the visual periphery, as a single summary statistic [20]. Critically, standard models of attention and working memory largely ignore ensemble representations, focusing instead on the representation of individual items [21–25]. Once ensembles and summary statistics are taken into consideration, it quickly becomes clear that observers have access to different aspects of the entire field of view, not just a handful of items.

In addition, we also discuss the idea that neural structures within the visual system involved in representing visual scenes and ensemble statistics [26,27] comprise a unique neural channel that is partially separate from other processing channels [28,29]. These results suggest that the visual system is functionally organized to allow for scene and ensemble representations to be efficiently formed somewhat independently of other object representations. In other words, there appear to be separate neural pathways for representing the forest and the trees.

Together, these findings help reconcile the apparent tension between our subjective impression of a rich visual world and empirical results highlighting the limits of visual cognition. We argue that the apparent richness of visual experience can be captured without having to dissociate consciousness from higher-level cognitive functions and without arguing that visual awareness overflows cognitive access.

The Limits of Visual Cognition

Two paradigms that have had a major role in demonstrating the limits of visual cognition are change blindness and inattentional blindness. Change blindness is the inability to detect a change between two different pictures when a brief interruption occurs between the two images [30,31] or the change occurs so gradually that it does not automatically draw attention [32]. By contrast, inattentional blindness is the failure to notice an otherwise visible stimulus when attention is directed elsewhere. In perhaps the most famous example, participants failed to notice a man in a gorilla costume walking through the middle of a scene when attention was focused on people passing a basketball [33]. Perhaps more commonly, automobile accidents regularly occur because drivers fail to notice items on the road (e.g., another car or a pedestrian) when their attention is directed elsewhere (e.g., their cell phone conversation) [34,35]. Despite differences in methodologies, both change blindness and inattentional blindness arise because of observers' limited ability to attend to and remember more than a few items at a time.

Although these paradigms clearly demonstrate the limits of visual cognition, more targeted studies have characterized the architecture and capacities of visual attention and working memory. Both of these processes are limited by a finite supply of some mental commodity [36]. This commodity is often characterized as either a fixed number of 'slots' [4,22–24] or a fluid cognitive resource [21,37,38]. Despite the differences between these models, they both converge on the idea that observers can store around three or four items in working memory. In terms of visual attention, initial studies estimated that around three or four locations can be attended at once [39,40], but more recent efforts have pushed that number closer to around seven or eight [25,41]. However, even eight attended locations is still not sufficient to explain the richness of perception.

Glossary

Attention: the process of selecting some bits of information for further processing at the expense of others (e.g., attending to the sound of a lecturer's voice and ignoring the street noise outside).
Awareness: the ability to consciously perceive, feel, or experience certain sensory events.
Bayesian inference: a method of statistical inference that uses Bayes' theorem to update the probability for a hypothesis as more information and/or evidence becomes available.
Ensembles and summary statistics: the representation of multiple items in the world as a single, average descriptor of the whole set (e.g., the average size of a collection of objects).
Gist of the scene: the basic perceptual (i.e., color, etc.) and conceptual (i.e., semantic label, etc.) representations of a scene that observers can comprehend in a single glance.
Recurrent processing: corticocortical interactions between neural regions in which information is transmitted from higher-level regions back to lower-level regions (e.g., from higher-level cortex back to early visual cortex).
Saccades: quick movements of the eyes that change the point of fixation.