The "item" as a window into how prior knowledge guides visual search

Rachel Wu[a] and Jiaying Zhao[b]

[a]Department of Psychology, University of California, Riverside, Riverside, CA 92521; [b]Department of Psychology and Institute for Resources, Environment and Sustainability, University of British Columbia, Vancouver, BC V6T 1Z4, Canada.
rachel.wu@ucr.edu   jiayingz@psych.ubc.ca

Abstract: We challenge the central idea proposed in Hulleman & Olivers (H&O) by arguing that the "item" is still useful for understanding visual search and for developing new theoretical frameworks. The "item" is a flexible unit that represents not only an individual object, but also a bundle of objects that are grouped based on prior knowledge. Uncovering how the "item" is represented based on prior knowledge is essential for advancing theories of visual search.

Hulleman & Olivers (H&O) present an elegant framework that aims to help us better understand visual search mechanisms. This framework proposes using fixations, rather than individual items, as the conceptual unit of visual search. The general ideas in the framework are very useful because it can account for many extant findings and identifies some shortcomings (such as embodied visual search) in the existing visual search literature.

Although this framework has its strengths, we disagree with the main argument that the item is no longer useful for understanding visual search. We do, however, agree with Olivers' earlier argument (Olivers et al. 2011) that visual search relies on an attentional template – a prioritized working memory representation – that is typically determined before starting a task via prior knowledge and/or explicit instructions. This attentional template evolves in various ways on the shorter time scale as the task progresses (e.g., Nako et al. 2015) and on the longer time scale as the learner gains more experience (e.g., Wu et al. 2015).

We argue that the "item" is still useful for understanding visual search and developing new theoretical frameworks. Critical to our argument is the idea that the "item" (contained in the attentional template) is a flexible unit that can represent not only an individual feature or object, but also a bundle of features or objects that are grouped based on prior knowledge. Such grouping, via either explicit or implicit cues, can result in the unitization of features or objects into an "item," which increases the amount of information held in working memory during visual search, and thus typically facilitates search performance. However, because many visual search studies control for prior experiences by using simple visual stimuli or equating prior knowledge across conditions, the nature and the limits of the attentional template are unclear. The use of prior knowledge is only mentioned briefly in H&O, but we believe that incorporating prior knowledge into visual search frameworks is critical for advancing the research area.

A growing number of studies on visual search (as well as visual working memory) demonstrate the benefits of prior knowledge on the outcomes of search tasks. For example, Nako et al. (2014a) confirmed that searching for one item (e.g., a letter) is more efficient than searching for two or more items (e.g., multiple letters), as evidenced by both neural measures (attenuated N2pc) and behavioral measures (slower reaction time and lower accuracy). Importantly, they demonstrated that if category knowledge can be applied during visual search, then one-item search and multiple-item search show very similar neural and behavioral outcomes. Nako et al. (2014b) and Wu et al. (2015) replicated and extended this initial finding using real-world objects, such as clothing, kitchen items, and human faces. In addition to prior knowledge about object category, grouping cues can also improve visual search. For example, Wu et al. (2016) showed that a heterogeneous set of novel alien stimuli grouped by an abstract rule (same versus different) can facilitate search performance.

Grouping of objects can occur not only by means of shared features and spatial proximity, but also by reliable co-occurrences over space and time. The visual system is remarkably efficient at detecting probabilities of co-occurrences among individual objects (e.g., Fiser & Aslin 2001; Turk-Browne et al. 2005), and this ability is present in early infancy (Fiser & Aslin 2002; Kirkham et al. 2002; Saffran et al. 1996; Wu et al. 2011). A direct consequence of learning the co-occurrences between objects is that the individual objects are implicitly represented as one unit (Mole & Zhao 2016; Schapiro et al. 2012; Wu et al. 2011; Wu et al. 2013; Zhao & Yu 2016). Such unitized representations implicitly and spontaneously draw attention to the co-occurring objects during visual search (Wu et al. 2013; Yu & Zhao 2015; Zhao & Luo 2014; Zhao et al. 2013), interferes with global processing of the visual array (Hall et al. 2015; Zhao et al. 2011), and increases the capacity of visual working memory (Brady et al. 2009; see also Brady et al. 2011). These findings support the idea that individual objects can be grouped into one "item" based on prior knowledge of co-occurrences, and such representations determine the allocation of attention, group objects into chunks, and facilitate search performance.

Besides the benefits of prior knowledge on visual search outcomes, there are also costs. When asked to search for one item in a category (e.g., the letter "A") and a foil item from the category appeared (e.g., the letter "R"), participants exhibited attentional capture to the foil at both neural and behavioral levels (Nako et al. 2014a). Wu et al. (2017) suggests that the "foil effect" is predicted by level of prior experience (e.g., distinguishing healthy and unhealthy foods based on dieting experience). Taken together, these recent studies show how the application of categorically based attentional templates (i.e., prior knowledge) can help overcome efficiency limitations in visual search by expanding the scope of target search, yet at the cost of false alarms to non-targets that fall within the search category.

In sum, we agree that investigating individual objects only may not provide a deeper understanding of visual search, but the "item" is still very useful. A better understanding of the bidirectional interactions of attention and learning allows us to build ecologically valid models reflecting cascading effects during visual search to advance this research area. Moreover, understanding how prior knowledge affects visual search and related attentional abilities has important implications for attention training. Given the growing literature showing the impact of knowledge on attention, increasing attentional abilities may involve training knowledge, rather than training attention per se.

Authors' Response

On the brink: The demise of the item in visual search moves closer

Johan Hulleman[a] and Christian N. L. Olivers[b]

[a]Division of Neuroscience and Experimental Psychology, School of Biological Sciences, The University of Manchester, Manchester M13 9PL, United Kingdom; [b]Department of Experimental and Applied Psychology, Institute for Brain & Behaviour Amsterdam, VU University, NL-1081 BT Amsterdam, The Netherlands.
johan.hulleman@manchester.ac.uk
https://www.research.manchester.ac.uk/portal/johan.hulleman.html
c.n.l.olivers@vu.nl
http://www.vupsy.nl/staff-members/christian-olivers/

Abstract: We proposed to abandon the item as conceptual unit in visual search and adopt a fixation-based framework instead. We treat various themes raised by our commentators, including the nature of the Functional Visual Field and existing similar ideas,