Cognition 126 (2013) 198–212

Contents lists available at SciVerse ScienceDirect

Cognition

journal homepage: www.elsevier.com/locate/COGNIT

The interplay of episodic and semantic memory in guiding repeated search in scenes

Melissa L.-H. Võ *, Jeremy M. Wolfe

Visual Attention Lab, Harvard Medical School, Brigham and Women's Hospital, USA

ARTICLE INFO

Article history:
Received 10 April 2012
Revised 21 September 2012
Accepted 30 September 2012
Available online 21 November 2012

Keywords:
Repeated and multiple search
Scene perception
Eye movements
Semantic memory
Episodic memory
Search guidance

ABSTRACT

It seems intuitive to think that previous exposure or interaction with an environment should make it easier to search through it and, no doubt, this is true in many real-world situations. However, in a recent study, we demonstrated that previous exposure to a scene does not necessarily speed search within that scene. For instance, when observers performed as many as 15 searches for different objects in the same, unchanging scene, the speed of search did not decrease much over the course of these multiple searches (Võ & Wolfe, 2012). Only when observers were asked to search for the same object again did search become considerably faster. We argued that our naturalistic scenes provided such strong "semantic" guidance—e.g., knowing that a faucet is usually located near a sink—that guidance by incidental episodic memory—having seen *that* faucet previously—was rendered less useful. Here, we directly manipulated the availability of semantic information provided by a scene. By monitoring observers' eye movements, we found a tight coupling of semantic and episodic memory guidance: Decreasing the availability of semantic information increases the use of episodic memory to guide search. These findings have broad implications regarding the use of memory during search in general and particularly during search in naturalistic scenes.

1. Introduction

We constantly interact with a complex environment that is predictable and variable at the same time. For example, we know that corkscrews generally rest on surfaces. They are usually found in kitchens rather than bedrooms and they most often inhabit the same drawer as the rest of the silverware does. Our knowledge of these regularities can be considered a form of semantic memory. These semantic regularities are probabilistic. Your search for the corkscrew could easily go astray at a friend's party if one of the guests deposited it in a low probability location. Under those circumstances, other cues would guide search. You might try to retrieve a vague memory of having seen the corkscrew on your friend's piano or you would go ahead and search for a small, shiny object with a corkscrew-like shape. When do we rely on guidance by probabilistic, semantic scene knowledge and when might we rely on episodic memory for a specific, previously noted location of that particular object?

In a recent study, we demonstrated that repeatedly searching for multiple different objects in the same, unchanging scene does not dramatically speed search despite the observer's increasing familiarity with the scene (Võ & Wolfe, 2012). Neither previewing nor memorizing each scene for 30 s produced marked benefits on subsequent object search. These results seem to run counter to our intuition that increased familiarity should improve the efficiency of search. In that case, we argued that search in our naturalistic scenes was guided by powerful scene semantics and that this strong semantic guidance minimized the usefulness of episodic memory in search guidance. In the present paper, we manipulate the availability

* Corresponding author. Address: Visual Attention Lab, Harvard Medical School, 64 Sidney Street, Suite 170, Cambridge, MA 02139, USA. Tel.: +1 617 768 8812.
E-mail address: mlvo@search.bwh.harvard.edu (M.L.-H. Võ).

0010-0277/$ - see front matter © 2012 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.cognition.2012.09.017

of semantic information in a scene in order to investigate the circumstances under which episodic memory will guide search.

1.1. Sources of guidance during search in naturalistic scenes

1.1.1. Feature guidance

From experiments using very simple displays, we know that there is a limited set of attributes that can be used to guide search. If you are looking for the large, red, tilted, moving vertical line, you can guide your attention toward the size, color, orientation, and motion of the items in a display. The idea of guidance by a limited set of basic attributes (somewhere between one and two dozen) can be called 'classic guided search' (see Wolfe, 2007; Wolfe & Horowitz, 2004). Search for the corkscrew would be aided by knowledge that it is shiny, has a very distinct shape, and is usually not bigger than your fist. Schmidt and Zelinsky (2009), for example, found that when targets were described using text, more descriptive cues led to faster searches than did less descriptive ones. More precise knowledge about *this* corkscrew's visual features, e.g. via a picture cue, would further speed search (Castelhano & Heaven, 2010; Castelhano, Pollatsek, & Cave, 2008; Malcolm & Henderson, 2009; Vickery, King, & Jiang, 2005). Similarly, Wolfe, Alvarez, Rosenholtz, and Kuzmova (2011, Exp.6) showed that when searching for an object twice in the same scene, search benefits were partially driven by learning to associate object specific features with the target word.

1.1.2. Semantic guidance

When targets are embedded in scenes, rather than in random arrays of items, search can draw on the rich information provided by the scene itself, in addition to any feature guidance. Over the course of our lifetime, we have learned to use the regularities encountered in our visual world to aid search. For instance, we learn to associate types of objects, e.g. any kind of toothbrush, to locations in certain types of scenes, e.g. a sink in any kind of bathroom scene. Thus, in addition to guidance by basic features, scenes offer "semantic" guidance, i.e. guidance by the structure and meaning of scenes. Semantic guidance allows drawing on a rich knowledge base—also referred to as sets of scene priors—readily accessible from even short glimpses of a scene (e.g., Castelhano & Henderson, 2007; Droll & Eckstein, 2008; Ehinger, Hidalgo-Sotelo, Torralba, & Oliva, 2009; Hidalgo-Sotelo, Oliva, & Torralba, 2005; Torralba, Oliva, Castelhano, & Henderson, 2006; Võ & Henderson, 2010). Semantic knowledge can be activated by the scene background and by specific, diagnostic objects in the scene. Diagnostic objects are those that by themselves strongly imply a certain scene category and/or the presence of other objects nearby. Thus, a toilet implies a bathroom and a table might imply nearby chairs. Semantic guidance, based on inter-object relationships within a scene, seems to be strong enough to guide search even when the background of a scenes is missing from a search display (see Wolfe et al., 2011, Exp. 5). The scene background provides its own information—like surface structures that objects might rest on—especially, when object-to-object relationships are weak. Thus, unlike a random display of isolated objects, a real scene itself can actually tell you where some objects are more likely to be found.

1.1.3. Episodic memory guidance

Contextual cueing studies have shown that even meaningless "scenes", in the form of repeated display configurations, can be learned in very short periods of time, with very simple items, and without observers' explicit awareness that they have been repeatedly exposed to the same target–distractor arrangements (Chun & Jiang, 1998; for a review see Chun & Turk-Browne, 2008). Observed benefits may result from more efficient allocation of attention to subsets of the visual display that most likely contains the target item or, perhaps, from enhanced decision processes (Kunar, Flusberg, Horowitz, & Wolfe, 2007). Classic contextual cueing paradigms provide evidence that the location of a particular target exemplar can be associated with a particular search array and thus might be taken as rather pure evidence that, even in the absence of semantic guidance, episodic memory for previous exposures to a scene can improve search.

When searching through real-world scenes, however, associations of targets to their context are often more abstract. For instance, we seem to be able to exploit relational contingencies that emerge across different scenes of the same category suggesting that statistical regularities abstracted across a range of stimuli are governed by semantic expectations (Brockmole & Võ, 2010). Further, search in naturalistic scenes seems to be biased to associate target locations to more global rather than local contexts (e.g., Brockmole, Castelhano, & Henderson, 2006; Brockmole & Henderson, 2006a; Brooks, Rasmussen, & Hollingworth, 2010). Unlike the usually implicit target-context associations in artificial displays, episodic memory for target-scene associations in real-world scenes tends to be explicit (Brockmole & Henderson, 2006b).

There is ample evidence that we have massive memory for objects (Brady, Konkle, Alvarez, & Oliva, 2008; Hollingworth, 2004; Konkle, Brady, Alvarez, & Oliva, 2010; Tatler & Melcher, 2007) as well as scenes (Konkle et al., 2010; Standing, 1973). Previously fixated (and therefore attended) objects embedded in scenes can be retained in visual long-term memory for hours or even days (e.g., Hollingworth, 2004; Hollingworth & Henderson, 2002; Hollingworth, Williams, & Henderson, 2001; for a review see Hollingworth, 2006). Even incidental fixations on objects during search improve subsequent recognition memory (e.g., Castelhano & Henderson, 2005; Hout & Goldinger, 2010; Võ, Schneider, & Matthias, 2008; Williams, Henderson, & Zacks, 2005). Moreover, Hollingworth (2009) showed that previewing a scene benefitted subsequent search through it. The effect of a preview increased with longer preview durations, consistent with evidence that observers accumulate visual scene information over the course of scene viewing (Hollingworth, 2004; Hollingworth & Henderson, 2002; Melcher, 2006; Tatler, Gilchrist, & Land, 2005; Tatler, Gilchrist, & Rusted, 2003).

To summarize, search in scenes can, in principle, benefit from a rich set of guiding sources: Feature, semantic, and episodic. Under which circumstances is one source of guidance prioritized over another?