Seek and you shall remember: Scene semantics interact with visual search to build better memories

Dejan Draschkow Ludwig-Maximilians University, Munich, Germany

Jeremy M. Wolfe Harvard Medical School, Cambridge, MA, USA
Brigham and Women's Hospital, Boston, MA, USA

Melissa L.-H. Võ Harvard Medical School, Cambridge, MA, USA
Brigham and Women's Hospital, Boston, MA, USA
Johann Wolfgang Goethe-Universität, Frankfurt, Germany

Memorizing critical objects and their locations is an essential part of everyday life. In the present study, incidental encoding of objects in naturalistic scenes during search was compared to explicit memorization of those scenes. To investigate if prior knowledge of scene structure influences these two types of encoding differently, we used meaningless arrays of objects as well as objects in real-world, semantically meaningful images. Surprisingly, when participants were asked to recall scenes, their memory performance was markedly better for searched objects than for objects they had explicitly tried to memorize, even though participants in the search condition were not explicitly asked to memorize objects. This finding held true even when objects were observed for an equal amount of time in both conditions. Critically, the recall benefit for searched over memorized objects in scenes was eliminated when objects were presented on uniform, non-scene backgrounds rather than in a full scene context. Thus, scene semantics not only help us search for objects in naturalistic scenes, but appear to produce a representation that supports our memory for those objects beyond intentional memorization.

Introduction

Imagine visiting a friend in his new home for the weekend. Arriving at his place, you might first search for a coat hook on which to hang your jacket. During your stay, you will certainly be looking for many things (the soap in the bathroom, the onions in the kitchen, or the printer in the home office, etc.). In preparation for your partner's inevitable questions regarding your friend's new home, you might also try to remember specific items you encountered (e.g., that Monet in the dining room). It seems evident that our memory for our visual surroundings will vary as a function of our interactions with the environment. Memorizing something explicitly (the Monet), or encoding information incidentally (the soap), should influence how many details we can subsequently recall from a scene. Further, the context in which objects are encoded should modulate memory for objects. In the three experiments presented here, we investigated memory recall as a function of task—memory vs. search—and as a function of context—semantically meaningful scenes vs. meaningless arrays of objects, to gain a better understanding of task-dependent scene representations.

Remembered scene representations can be very detailed and can endure for long periods of time (Hollingworth & Henderson, 2002; Hollingworth, 2004, 2006; Konkle, Brady, Alvarez, & Oliva, 2010). Under the right circumstances we have massive memory for individual objects as well as scenes (Brady, Konkle, Alvarez, & Oliva, 2008; Hollingworth, 2004; Konkle et al., 2010; Standing, 1973; Tatler & Melcher, 2007). Within scenes, memory performance is predicted by gaze durations (Hollingworth & Henderson, 2002) and number of fixations (Tatler, Gilchrist, & Land, 2005; Tatler & Tatler, 2013) on critical objects. Objects that are task-relevant are remembered better than task irrelevant ones (Castelhano & Henderson, 2005; Maxcey-Richard & Hollingworth, 2013; Williams, Henderson, & Zacks, 2005), but even the irrelevant ones are reliably incidentally encoded (Castelhano & Henderson, 2005; Hollingworth, 2006; Võ, Schneider, & Matthias, 2008). Of course, scene memory is not

Draschkow, D., Wolfe, J. M., & Võ, M. L.-H. (2014). Seek and you shall remember: Scene semantics interact with visual search to build better memories. *Journal of Vision*, *14*(8):10, 1–18, http://www.journalofvision.org/content/14/8/10, doi:10.1167/14.8.10.

doi: 10.1167/14.8.10 Received January 30, 2014; published July 11, 2014 ISSN 1534-7362 © 2014 ARVO

perfect. For example, boundary extension is a memory error often assessed with explicit drawing tasks. First demonstrated by Intraub and Richardson (1989), boundary extension produces evidence for "memory" for unseen spaces, inferred beyond the boundaries of the view actually presented. After being shown scenes, people often remember seeing parts of the scene that were never visually presented. Intraub (2012) suggested that scene-memory is not limited to visual representation, but rather a picture is understood to be part of a continuous world and not simply as a piece of paper with shapes and colors on it.

Nevertheless, those findings suggest that durable representations of naturalistic scenes are formed during a variety of tasks. In our initial example, memorizing the Monet in the dining room for later recall should be a feasible task, as you explicitly allocated memory resources to accomplish this goal. But how about objects (such as the soap in the bathroom) that you might have searched for, but never intended to memorize? Obviously, you learn about the layout of the world without explicitly instructing yourself to remember, but how does this implicit encoding through the active engagement with an object compare to its intentional encoding under explicit instruction?

Memory in search through displays of simple stimuli randomly placed on uniform backgrounds is usually inferred by reaction times (RTs). In contextual cueing, for example, memory of the layout of distractors on earlier trials speeds RT on subsequent trials when that layout is repeated (Chun & Jiang, 1998). Several paradigms have been used to provide evidence for memory within and/or across trials for incidentally encountered items (e.g., Boot, McCarley, Kramer, & Peterson, 2004; Hout & Goldinger, 2010, 2012; Howard, Pharaon, Körner, Smith, & Gilchrist, 2011; Klein & MacInnes, 1999; Körner & Gilchrist, 2007, 2008; Kristjánsson, 2000; Peterson, Beck, & Vomela, 2007; Peterson, Kramer, Wang, Irwin, & McCarley, 2001; Solman & Smilek, 2010). On the other hand, there are also situations in which participants do not appear to make use of memory for previously encountered distractors (e.g., Horowitz & Wolfe, 1998, 2003). In some cases, incidentally encountered items may be remembered, but that memory might not be used in a subsequent search because it is simply faster to search de novo than to retrieve specific memory (Kunar, Flusberg, & Wolfe, 2008). However, RTs alone cannot tell us whether participants really remember encountered stimuli but prefer not to use this memory, or whether they simply cannot recall what they saw.

When searching through more naturalistic scenes, additional guidance factors need to be considered (Tatler, Hayhoe, Land, & Ballard, 2011; Wolfe, Võ, Evans, & Greene, 2011). By any estimate of the "set size" in a natural scene, searches for objects in scenes seem to be very efficient (Wolfe, Alvarez, Rosenholtz, Kuzmova, & Sherman, 2011; Wolfe et al., 2010) compared to searches for isolated objects (Vickery, King, & Jiang, 2005). This difference is due, at least in part, to the availability of "semantic guidance," i.e., guidance by the structure and meaning of scenes. Semantic guidance is based on a rich knowledge base of what can be called "scene priors" (e.g., knowing that the jam jar is not likely to be in the bathtub). Much of this knowledge can be activated by glimpses of a scene as short as 50 ms and subsequently used to efficiently guide search (Võ & Henderson, 2010). These priors can be considered a form of memory. Unlike mainly implicit phenomena like contextual cueing (Chun & Jiang, 1998; Chun, 2000), much of the memory for more naturalistic scenes is explicit (Brockmole & Henderson, 2006; Hollingworth & Henderson, 2002). Thus, scene semantics not only guide search, but scene semantics seem to also support memory as can be seen in "massive memory" for objects and scenes (Brady et al., 2008; Konkle et al., 2010).

While it seems unquestionable that we are able to extract and retain scene information in episodic scene memory, the degree to which episodic scene memory is actually used to guide search can be debated. Võ and Wolfe (2012), for instance, demonstrated the dominance of guidance by generic scene knowledge over specific, episodic scene memory in real-world search. In a series of experiments, familiarity with a scene, derived from exposure or explicit memorization, did not necessarily benefit subsequent search for other objects embedded in that scene. The authors argued that search in naturalistic scenes was so effectively guided by knowledge of the semantics and syntax of scenes (i.e., *what* objects should be *where* within a scene, see Võ and Wolfe, 2013a), that the usefulness of episodic memory in scene search was minimized. This seems to also be the case in more naturalistic, 3D environments, where participants search by actually moving their bodies in space (Kit et al., 2014; but see Hollingworth, 2012 for contrary arguments). However, when the same target object was searched for a second time many trials later, memory for the previous search significantly speeded search. Recently, Võ and Wolfe (2013b) used 3D rendered images of real-world scenes to investigate the interplay of semantic and episodic memory guidance during search in naturalistic scenes. They found that decreasing the availability of semantic information by placing objects in unexpected locations increased the use of episodic memory in guiding search, demonstrating a complex interplay of semantic and episodic memory guidance during real-world search.

Note that these studies never actually tested scene memory explicitly, but inferred use of some sort of memory from the effects of memory on RT. This is an indirect measure of memory and does not provide us with information about the extent to which participants were able to establish scene representations as a function of encoding tasks.

A recent study investigated how memory performance depended on encoding instructions in a real world setting (Tatler & Tatler, 2013). Eye movements were recorded with a portable eye tracker, while participants were exposed to three different conditions. In the free viewing condition, participants walked around in a room without specific instructions and were therefore able to adopt their own task and strategy. In the undirected memory condition, participants had to remember as much as possible about every object in the room. Finally, in the directed memory condition, they were asked to focus on only a subset of objects (here the ones needed to make tea). In this case, the authors assessed memory performance directly with a 4AFC questionnaire and, thus, could make inferences about the impact of instructions on subsequent recall. Results showed that performance was above chance in the free viewing task, which lacked intentional encoding, and was presumably based on incidental encoding (see Castelhano & Henderson, 2005; Williams et al., 2005). However, memory for objects was much better in the memory conditions. In particular, relevant objects (the tea making objects) were remembered better than irrelevant ones. Based on the eye movement data, Tatler and Tatler argued that instructions modulated memory performance via a combination of strategic differences in fixation allocation and strategic differences in the extraction and retention of information from fixations. Thus, the task at the time of encoding modulates to what degree information about a scene is going to be memorized.

In the present study we contrast intentional memorization with incidental encoding during search using recall as our measure of memory. Based on Tatler and Tatler's (2013) findings, one might assume that explicit instructions to memorize a set of critical objects would create stronger memory representations than mere incidental encoding. However, that experiment used free viewing as its incidental memory condition. As noted above, other studies suggest strong incidental memory for objects that were the targets of visual search (e.g., Hollingworth, 2012; Võ & Wolfe, 2012, 2013b; Wolfe et al., 2011). The goal of the present study was to directly compare memory representations generated implicitly during search to those created by explicit memorization instructions. As an explicit recall task, we asked our participants to draw the scenes from memory, since this type of explicit recall test assesses both local and global characteristics of stored scene representations. In Experiments 1 and 2, participants either searched for target objects embedded in repeating scenes or were asked to memorize designated objects in other scenes. Afterwards we instructed participants to draw the scenes they had searched or memorized. We also tracked participants' eye movements while they searched and memorized scenes to compare gaze durations on objects between the two tasks and evaluate their contribution to differences in free recall performance. Interestingly, we found that despite the lack of explicit instructions to memorize objects in scenes, searched objects showed superior memory performance compared to memorized objects. In Experiment 3, we examined the role scene semantics play in the generation and retrieval of scene memories. Participants searched for or explicitly memorized isolated objects that were superimposed on distinctive backgrounds, but crucially, the objects were not embedded in scenes that would provide semantic support for memory. We found that without semantic guidance, the memory performance benefit for searched objects disappeared. Indeed, now participants recalled less searched than explicitly memorized objects. Between experiment comparisons indicated that this was due to a performance drop for searched items, rather than an increase for memorized ones.

General methods

Participants

In each of the three experiments, 10 participants were tested (Experiment 1: Mean Age = 28.6, $SD = 6$, 4 female; Experiment 2: M = 21.5, $SD = 3.5$, 7 female; Experiment 3: M = 30.7, $SD = 9.4$, 8 female). All were paid volunteers who had given informed consent. Each had at least 20/25 visual acuity and normal color vision as assessed by the Ishihara test.

Stimulus material

Ten full-color images of indoor scenes were used in Experiments 1 and 2 (for an example of the images see Figure 1, left column). An additional image was used for practice trials. Images were carefully chosen to include 15 singleton targets; that is, when searching for a wine glass only one object resembling a wine glass would be present in the scene. In Experiment 3, 15 such objects were displayed randomly on uniform, but distinct backgrounds (for an example of the images see Figure 1, right column). These objects were chosen to resemble the objects used in the scenes from Experiments 1 and 2. Thus, the collections of objects presented on a uniform background in Experiment 3 were matched to the objects populating the original scenes in that all the "office" objects from an

Figure 1. Examples of the stimulus material used in the study. Images of indoor scenes used in Experiments 1 and 2 (left column). Experiment 3 used ten full-color images of 15 objects displayed randomly on essentially uniform, but distinct backgrounds (right column). The objects were chosen to resemble the objects used in the scenes from Experiments 1 and 2, i.e., if a "monitor" was used in the "office" scene a similar "monitor" was used on the uniform background (compare upper left and right images).

Experiment 1 scene would be placed on one uniform background in Experiment 3, albeit in random locations. Scenes were displayed on a 19-inch computer screen (resolution 1024 × 768 pixel, 100 Hz) subtending visual angles of 37° (horizontal) and 30° (vertical) at a viewing distance of 65 cm.

Apparatus

Eye movements were recorded with an EyeLink1000 desktop system (SR Research, Canada) at a sampling rate of 1000 Hz. Viewing was binocular but only the position of one eye was tracked. Experimental sessions were carried out on a computer running OS Windows XP. Stimulus presentation and response recording were controlled by Experiment Builder (SR Research, Canada).

Procedure

Each experiment was preceded by a randomized 9-point calibration and validation procedure. After every tenth trial, a drift check was applied. If necessary, a recalibration was performed.

In all experiments, participants were presented with scenes in two subsequent conditions (Search and Memory); In the *Search condition,* participants searched for 10 different objects in each of five unchanging scenes. The search target was identified at the start of each search by presenting a target word (font: Courier; font size: 25; color: white with black borders; duration: 750 ms) at the center of the scene. Participants were instructed to search for the object as fast as possible and, once found, to press a button while fixating the object. From the point the target word appeared until a response was made the scene remained

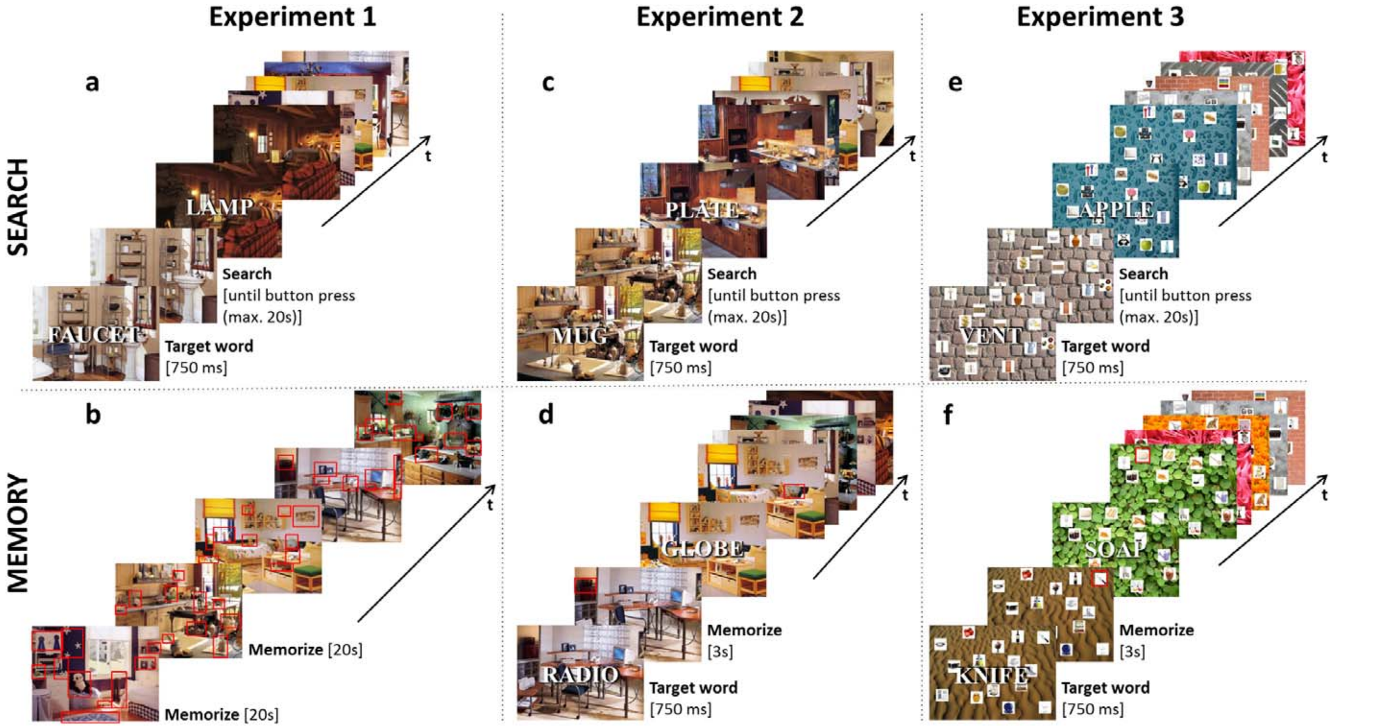

Figure 2. The trial procedure for the Search condition in Experiment 1 (a), Experiment 2 (c) and Experiment 3 (e), and the Memory condition in Experiment 1 (b), Experiment 2 (d) and Experiment 3 (f). Target words at the center of the display designated the target object of each search (a, c and e), or the critical object (red frame) for each memorization (d and f).

continuously visible. The next search started with the appearance of a new target word (see Figure 2a, 2c, and 2e).

The Search condition consisted of 50 trials (10 searches in each of five scenes) with scenes and objects in random order; i.e., after searching for a shaving brush in a bathroom, the next search could be for a different object within the same or in one of the four other scenes. Each object was searched for only once. The procedure for the Search condition was the same for all three experiments.

In the *Memory condition,* participants were instructed to memorize as much as possible of each of five scenes for a later memory test. Specifics of this condition were varied between experiments. In Experiment 1, the five scenes were presented consecutively for 20 seconds each and the 10 critical objects per scene were simultaneously surrounded by red frames. Participants were instructed to memorize as much as possible, but they were explicitly told to prioritize memorization of the 10 outlined objects (see Figure 2b). The instruction emphasized memorization of marked target objects, but also mentioned the scene background, as a memorization task in which participants focus only on the framed objects and disregard the scene context would make the two tasks more dissimilar. After the presentation of each scene, a drift check was performed for eye tracking. In Experiments 2 and 3, the procedure was designed to closely mimic the Search condition but without the search component. The object to be memorized was identified at the start of each trial by presenting a target word for 750 ms at the center of the scene. At this point, the critical object was immediately surrounded by a red frame to eliminate the need for a search. Participants were instructed to prioritize memorization of the object but to memorize the scene as well. After 3 s of memorization time, the next trial started with the appearance of a new target word and framed object (see Figure 2d and 2f).

Task order (Search vs. Memory) and scene assignments were counterbalanced across participants. The order of object searches was Latin square randomized such that across participants each of the 150 objects was equally often a target. More detailed descriptions of these tasks can be found in the specific methods sections for each experiment. An additional scene was used for practice trials at the beginning of each experiment, which was not included in the final analyses.

After each of the two conditions, participants were asked to draw all five scenes with as many objects as possible from memory on prepared evaluation sheets shown in the Supplementary Material. For this recall memory task, participants received one sheet at a time and were asked to draw everything they could remember. They were allowed to place labels of objects in the drawing of the scene if they felt uncomfortable drawing the objects. In order not to suggest or suppress any type of strategy, participants could choose the sequence of scenes they would draw. The only requirement for the participants was that once they finished a scene, they were not allowed to go back to it. They were also told to consider the edges of the rectangle on the sheet to be the edges of the scenes, filling in the space as it had been filled in the photograph (see Intraub & Richardson, 1989). No other limitations or temporal constraints were set, in order to allow a maximum of freedom for participants to access their memory of the scenes. Whereas participants in the Memory condition had been explicitly told about the upcoming memory test, they had been left unaware of this in the Search condition.

After participants finished the recall task, we also assessed the possibility of boundary extension using a modified camera distance paradigm (e.g., Gagnier & Intraub, 2012; Intraub & Richardson, 1989). We again presented the 10 scenes participants had encountered in the two conditions (five per task). The scenes were shown for 3 s each, and after every image participants were asked to indicate on a 4-point rating scale whether the test picture was *a lot closer up* (1), *a little closer up* (2), *a little further away* (3), or *a lot further away* (4) than it had been when they had to search through/memorize it. If participants indicated that the test picture seemed closer up, this response would imply that they had remembered scene details extending the boundaries of the presented one making the currently presented scene seem closer up. On the contrary, choosing the alternatives (3) or (4) suggests that less of the scene was remembered, implying the restriction of boundaries in memory. Note, that participants could not choose the only right answer ("identical").

Each experiment lasted between 40 min and 1 hr.

Data analysis

Raw eye movement data were filtered using SR Research Data Viewer. As already mentioned, 15 objects were preselected as critical in each scene. Targets were defined as the 10 critical objects per scene a participant had to search for or memorize. Due to the vast set size of objects in naturalistic images, five random objects per scene served as control objects, in order to have an arbitrary measure of recall performance for objects that were not explicitly relevant to the task. These distractors were never relevant to either the search or explicit memorization task. This was done to investigate target-distractor recall differences as an indicator of task relevance manipulations. Objects were picked via a random draw, so that no participant had the same combination of target and distractor objects. The interest area for each target/distractor object was defined by a rectangular box that was large enough to encompass that object. A search was deemed successful when the participants pressed the button while fixating the target's interest area or when that interest area was fixated within 500 ms prior to or after the button press. Summed gaze durations for each object were calculated by summing up the time spent fixating an object's interest area throughout all presentations of the scene both before and after an object became a search/memory target. Trials in which gaze duration on an object exceeded 7.5 s were excluded as outliers (in all experiments <1 %). Gaze durations of this length would either indicate that participants were not doing the task or technical issues.

To assess memory performance, we calculated the mean percentages of targets and distractors that were drawn by the participant as well as the mean percentage of recalled objects that were accurately located with respect to the rest of the scene. Location accuracy was coarsely defined by dividing the scene into quadrants (relative to each participant's specific perspective of the drawn scene) and coding whether the recalled object was also drawn in the correct quadrant or not (see Supplementary Material). The most important criteria for the accuracy of a drawing were the presence or absence of an object. A recalled object which was not recognizable or meaningfully labeled was not counted as correctly recalled. To avoid experimenter bias, two experimenters independently assessed recall performance and in case of conflict/uncertainty, a third experimenter was added for adjudication.

We used recognition memory for boundaries (see Intraub & Richardson, 1989) as an additional memory performance measure that included the scene as a whole. To analyze boundary extension, we looked at participants' responses to the final recognition task. Perfect boundary memory was defined as the mean of response possibilities ($=2.5$). Lower values (<2.5) would indicate boundary extension, whereas higher values (>2.5) stand for boundary restriction.

We conducted planned comparisons of gaze durations and memory performance between the two conditions (Search vs. Memory) using Welch's t tests separately for targets and distractors. To assess if task-relevant objects were recalled more often, we also compared performance for targets and distractors within each task. To see if gaze durations were similarly modulated by task, we conducted an ANOVA with the within-subject factors Targets Recalled (Present vs. Absent) and Task (Search vs. Memory) for dwell times.

We also tested the effect of scene semantics by comparing memory performance and gaze durations between Experiment 2 and 3 using between-subject comparisons (Experiment 2: Present vs. Experiment 3: Absent) for both Search and Memory conditions.

Experiment 1

In this first experiment, we investigated the explicit memory representation of real-world scenes after incidental encoding during object search compared to the representations of intentionally memorized scenes. We tracked participants' eye movements while they searched or memorized scenes to compare the amount of time participants looked at objects in the corresponding tasks. Previous work showed that searched objects produced faster responses than other objects (Hollingworth, 2012; Võ & Wolfe, 2013b). However, as noted, these studies used indirect measures of recall. Experiment 1 directly addresses the lack of an explicit memory test in the previous studies by asking participants to redraw previously presented scenes. In search and memorization conditions, we assessed subsequent recall for a critical set of objects in the scene.

Procedure

As described in the General methods section, participants were asked to search for 10 different objects within each of five unchanging scenes in the Search condition (Figure 2a). In the Memory condition, the same participants were instructed to memorize five consecutively presented scenes for 20 seconds each, putting special emphasis on the 10 objects that were simultaneously surrounded by red rectangular frames (Figure 2b). After each of the conditions, recall performance was assessed by asking participants to draw each of the scenes with as many of their objects as possible. Participants were unaware of the upcoming memory test while performing in the Search condition. They were aware of the memory test in the Memory condition.

Results

Objects recalled

Figure 3 depicts the memory performance for objects (both targets and distractors separately) and their locations. Recall performance was assessed by computing the probability that an object was drawn as a function of the task condition (Memory vs. Search). We present recall performance separately for task-relevant objects (targets) and irrelevant ones (distractors). Participants recalled twice as many targets after the Search condition (42%), than the Memory condition (20%), $t(9) = 6.72$, $p < 0.01$ (Figure 3a). There was no difference for remembered distractors between tasks, $t(9) < 1$ (Figure 3b). The relevance of objects had a major impact on recall performance and was modulated by task: The percentage of targets recalled was almost six times greater than the percentage of distractors in the Search condition, $t(9) = 8.92$, $p < 0.01$, while this relevance effect was observable, but smaller after memorization, $t(9) = 2.92$, $p < 0.02$. There was a significant difference in the magnitude of this effect between tasks, $t(9) = 5.61$, $p < 0.01$, pointing towards higher values during search (Search: M = 34.8%, SEM = 3.9% versus Memory: M = 11.4%, SEM = 3.9%).

Location accuracy

Only locations of drawn objects were assessed. Overall a high percentage (81%) of objects was drawn at accurate positions. Targets were positioned equally accurately in both conditions, $t(9) < 1$ (Figure 3c). Two participants were excluded from analysis on distractor location accuracy, as they did not recall any distractor objects in the Search condition. Distractors were positioned equally accurately in both conditions, $t(7) = 1.47$, $p > 0.1$ (Figure 3d).

Boundary extension

Participants were surprisingly good in remembering image boundaries. There was neither boundary extension for Search, $t(9) < 1$, nor Memory conditions, $t(9) < 1$. Mean boundary extension values did not differ between conditions, $t(9) = 1.34$, $p > 0.2$.

Task order effects on recall performance

There was no main effect of Task order on recall performance, $F(8, 1) < 1$ and no interaction between Task and Task order, $F(8, 1) < 1$. Only the main effect of Task was significant, $F(8, 1) = 40.55$, $p < 0.01$. The order of the two tasks neither affected recall in the Search condition, $t(7.89) < 1$, nor in the Memory condition, $t(7.35) < 1$.

Gaze durations

In Figure 4 we show the mean gaze durations on objects during search and memorization. There was a trend towards longer gaze durations for target objects during search as compared to memory, $t(9) = 2.15$, $p = 0.06$ (Figure 4a). Distractors were looked at for significantly longer during the Search condition, $t(9) = 3.06$, $p < 0.02$ (Figure 4b). Within the Memory condition, participants gazed longer on targets, $t(9) = 8.70$, $p < 0.01$. The same was true during search, $t(9) = 9.28$, $p < 0.01$.

Dwell times differed significantly depending on whether a target was recalled or not, $F(9, 1) = 16.87$, $p < 0.01$, showing that recalled targets received higher

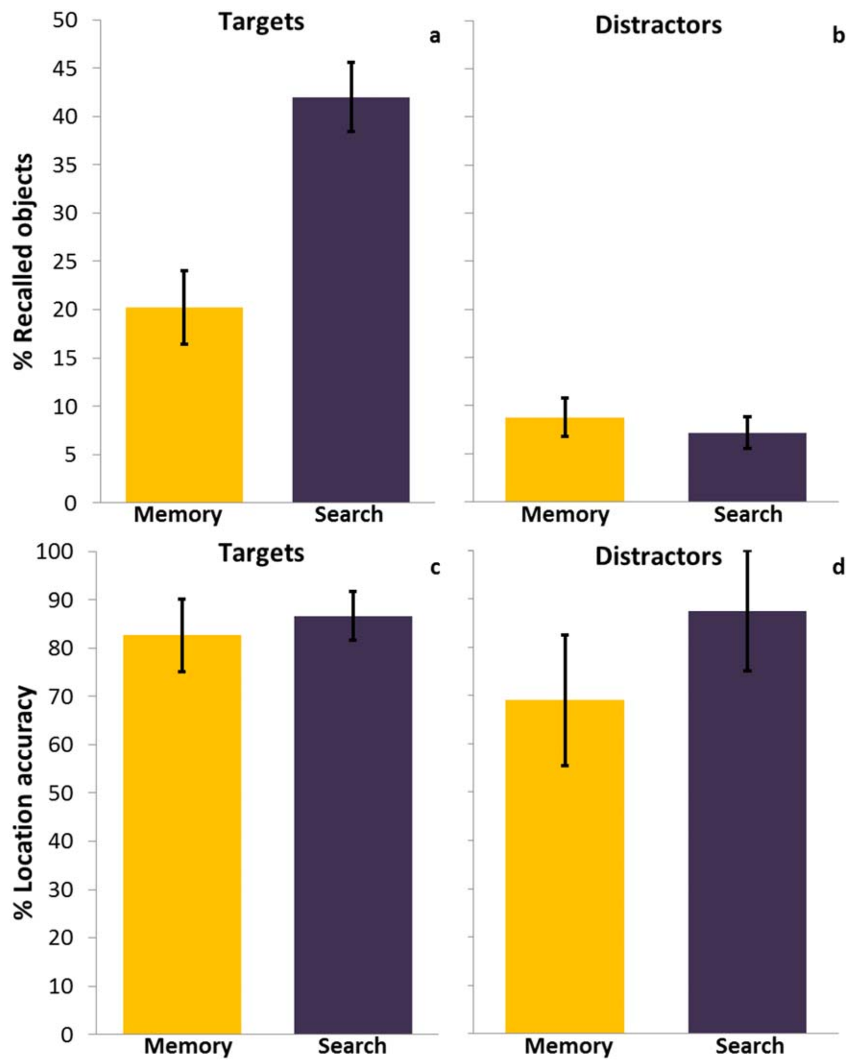

Figure 3. Memory performance as assessed by the mean percentage of recalled objects (a and b) and location accuracy for these recalled objects (c and d) depicted separately for targets (a and c) and distractors (b and d) as a function of Task (Memory=yellow vs. Search=purple).

gaze durations than targets that were not drawn (Figure 4c). There was no main effect of Task, $F(9, 1) = 4.67$, $p = 0.06$ and no significant interaction, $F(9, 1) = 1.34$, $p > 0.2$. Because two participants did not recall any distractor objects in the Search condition, they were excluded from distractor analysis. The remaining participants did not gaze longer on distractors which were subsequently recalled compared to ones which were not recalled, $F(7, 1) = 2.57$, $p > 0.1$. There was no main effect of Task, $F(7, 1) = 5.46$, $p > 0.05$ and no significant interaction, $F(7, 1) < 1$.

Discussion

Replicating previous results (Castelhano & Henderson, 2005; Tatler & Tatler, 2013; Võ et al., 2008; Williams et al., 2005), we found that task-relevant objects in both conditions produced a stronger memory representation, than task irrelevant ones. This effect was stronger, when participants had to search for, rather than memorize objects. Nevertheless some distractors were also recalled, a result which also falls in line with previous findings.

The key finding is that instructions to search for specific objects produced higher rates of recall than the explicit request to memorize them. Participants recalled twice as many targets after search than after memorization (42% against 20%) (Figure 3). This is surprising since participants were unaware of a subsequent recall task in the search condition.

The interpretation of the search-superiority finding is complicated by the fact that participants looked for a slightly longer period of time at searched targets than at memorized ones (Figure 4). The longer viewing times might therefore have accounted for the search superi-

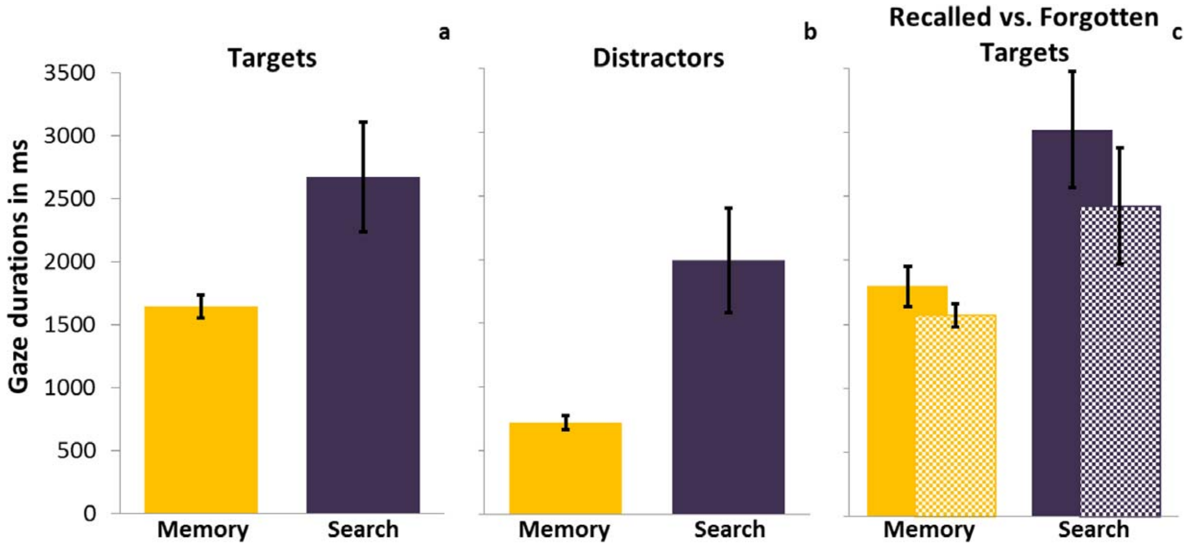

Figure 4. Gaze durations as a function of Task (Memory=yellow vs. Search= purple) for target objects (a) and distractors (b). The right graph (c) depicts gaze durations on targets as a function of Task (Memory=yellow vs. Search= purple) and Targets Recalled (Present=full vs. Absent=checkered).

ority. Consequently, in Experiment 2, we tried to better match both gaze durations and trial sequences for Search and Memory conditions.

Experiment 2

Although it is impossible to perfectly equate the experience of searching with the experience of memorization, we tried to make both tasks as similar as possible. Thus, in Experiment 2, we tried to match gaze durations between the two tasks by employing longer scene presentations in the Memory condition (>30 s compared to 20 s in Experiment 1). In addition, the simultaneous presentation of all framed target objects during memorization in Experiment 1 might have impaired encoding performance. Perhaps, participants distributed their attentional resources inefficiently when faced with ten critical objects at once. This or some other difference in the presentation of objects in the Memorization versus Search conditions might have produced fewer objects in recall. To control for these possible issues, the procedure in the Memory condition of Experiment 2 was designed to closely mimic the Search condition, while leaving out the search component.

Procedure

In order to make the Memory condition as similar as possible to the Search condition, the object to be memorized was identified at the start of each trial by presenting a word label for 750 ms in the center of the scene. The critical object was then immediately surrounded by a red frame (for 3 s), thus eliminating the need for any act of search since the salient frame attracted attention immediately (see Figure 2d). Participants were instructed to memorize the object in the context of the broader scene. This produces 30 s of explicit memorization time and 37.5 s of total exposure to the scene. On each of the 50 trials, a new target had to be memorized. The appearance of a new target word initialized the subsequent trial. Presentation of target objects and scenes was randomized with all five scenes interlaced with each other. The search task condition was performed as in Experiment 1.

Results

As in Experiment 1, we present memory performance and gaze durations separately for task-relevant (targets) and irrelevant objects (distractors).

Objects recalled

Replicating the findings from the first experiment, target recall was better after search (45%) than after explicit memorization (32%), $t(9) = 2.44$, $p < 0.05$ (Figure 5a). Again, there was no difference for remembered distractors between tasks, $t(9) < 1$ (Figure 5b). Greater percentages of targets than distractors were reproduced in both the Search, $t(9) = 7.79$, $p < 0.01$ and Memory condition, $t(9) = 7.31$, $p < 0.01$. The magnitude of this effect was larger in the Search condition, $t(9) = 3.14$, $p < 0.02$ (Search: M = 35.8%, *SEM* = 4.6% versus Memory: M = 22%, *SEM* = 3%).

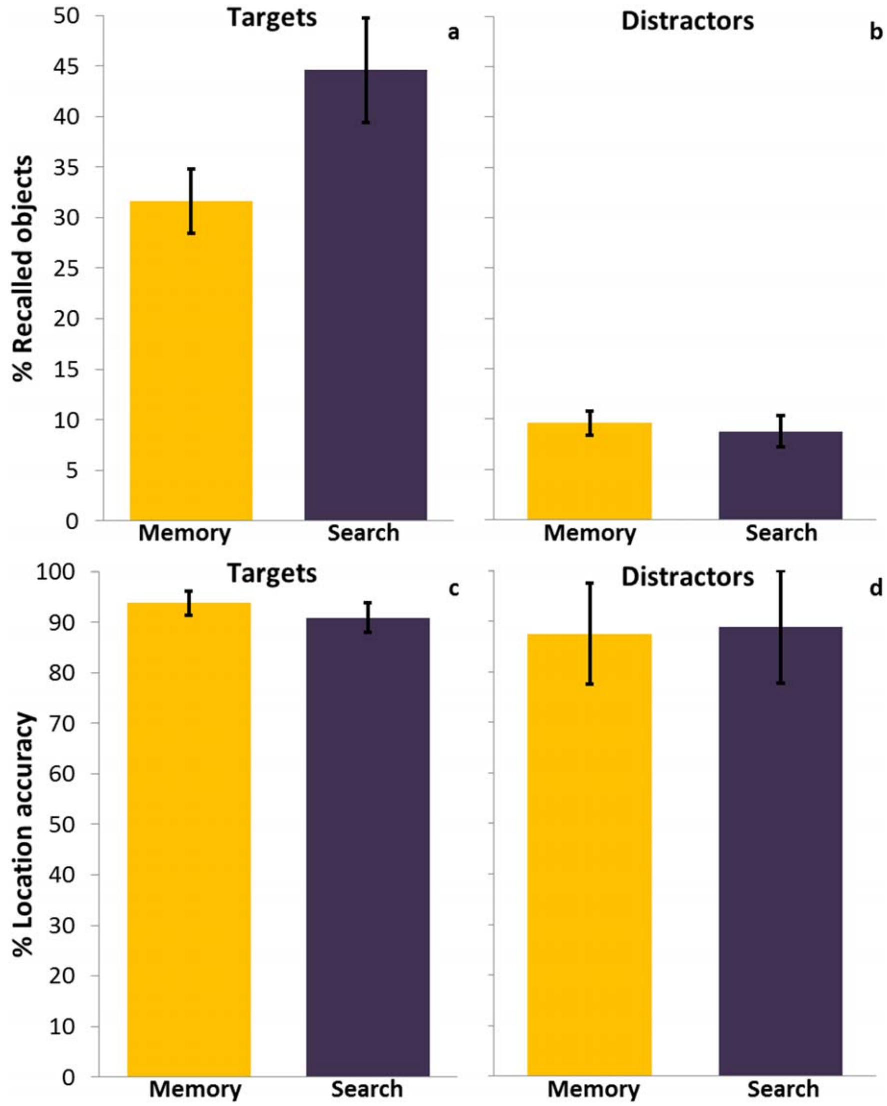

Figure 5. Memory performance as assessed by the mean percentage of recalled objects (a and b) and location accuracy for these recalled objects (c and d) depicted separately for targets (a and c) and distractors (b and d) as a function of Task (Memory=yellow vs. Search= purple).

Location accuracy

High percentages (90%) of recalled objects were drawn in roughly accurate positions. Targets were positioned equally accurately in both conditions, $t(9) < 1$ (Figure 5c). One participant was excluded from the analysis on distractor location accuracy, as he did not recall any distractor objects in the Search condition. Distractors were positioned equally accurately in both conditions, $t(8) < 1$ (Figure 5d).

Boundary extension

Data from one participant was not collected, because of technical issues with the boundary extension program. As in Experiment 1, participants were surprisingly good in remembering image boundaries. There was no evidence for boundary extension in Search, $t(8) < 1$, or Memory conditions, $t(8) < 1$. Mean boundary extension values did not differ between conditions, $t(8) < 1$.

Task order effects on recall performance

There was no main effect of Task order on recall performance, $F(8, 1) < 1$ and no interaction between Task and Task order, $F(8, 1) = 1.22$, $p > 0.05$. The main effect of Task did not become significant, $F(8, 1) = 1.06$, $p > 0.05$. The order of the two tasks neither affected recall in the Search condition, $t(7.97) < 1$, nor in the Memory condition, $t(7.92) < 1$.

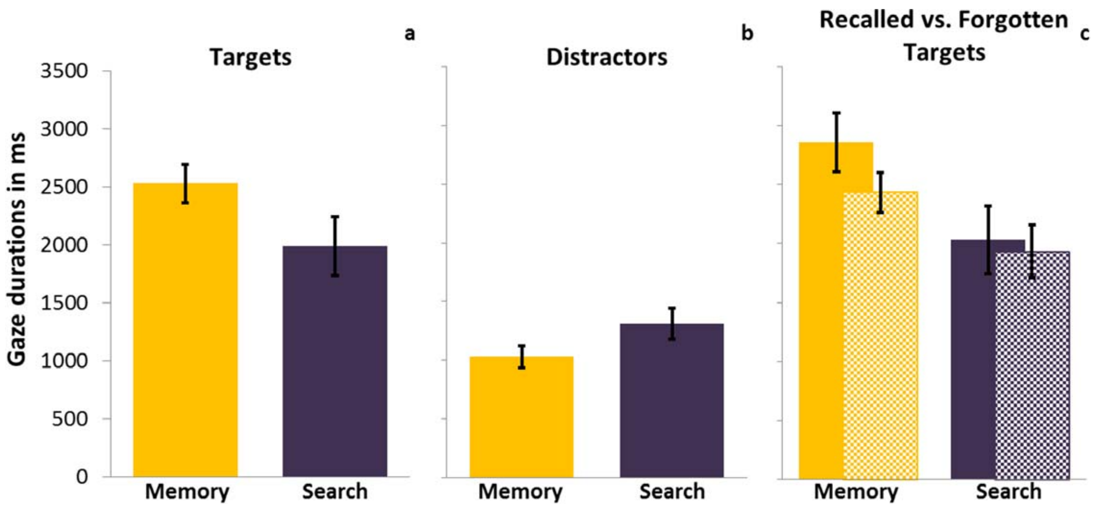

Figure 6. Gaze durations as a function of Task (Memory=yellow vs. Search= purple) for target objects (a) and distractors (b). The right graph (c) depicts gaze durations on targets as a function of Task (Memory=yellow vs. Search= purple) and Targets Recalled (Present=full vs. Absent=checkered).

Gaze durations

Figure 6 depicts the mean gaze durations on objects during search and memorization. The goal of Experiment 2 was to equalize the exposure to targets in the Search and Memory conditions. In contrast to Experiment 1, gaze durations on targets were now slightly longer in the Memory condition compared to the Search condition, though this difference was significant neither for targets, $t(9) = 1.73$, $p > 0.1$ nor distractors, $t(9) = 1.63$, $p > 0.1$ (Figure 6a and 6b). In any case, this result means that the better recall for search tasks cannot be simply attributed to greater exposure in the Search condition of this experiment. Participants gazed longer on targets than distractors in both Search, $t(9) = 4.28$, $p < 0.01$ and Memory conditions, $t(9) = 8.39$, $p < 0.01$. The main effect of Targets Recalled on dwell times did not reach significance, $F(9, 1) = 5.01$, $p > 0.05$ and neither did the main effect of Task, $F(9, 1) = 4.24$, $p > 0.05$ nor the interaction, $F(9, 1) = 2.25$, $p > 0.1$ (Figure 6c). One participant did not recall any distractor objects in the Search condition, thus was excluded from distractor analysis. Dwell times differed significantly depending on whether a distractor was recalled or not, $F(8, 1) = 1.12$, $p < 0.05$, showing that recalled distractors received higher gaze durations than distractor objects that were not drawn. There was no main effect of Task, $F(8, 1) = 2.15$, $p > 0.05$ and no significant interaction, $F(8, 1) = 2.00$, $p > 0.05$.

Discussion

By designing the Memory condition to mimic the Search condition as closely as possible in Experiment 2, we roughly matched gaze durations between conditions. If anything, gaze durations for objects in the Search condition are now slightly shorter than in Memory condition (Figure 6). Nevertheless, we replicated the recall memory performance results from Experiment 1: Despite the lack of explicit instructions to memorize objects in scenes, searched objects were retrieved better than memorized objects (Figure 5). This preference suggests that the similar pattern of results in Experiment 1 was not an artifact of the time each object was fixated by the participant.

Experiment 3

In Experiment 3, we wanted to investigate if semantic information in natural scenes played a critical role in the counterintuitive superiority of recall for searched over memorized targets. In Experiments 1 and 2, participants searched for/memorized objects in real scenes. Would the search benefit remain if the targets were not part of a meaningful scene? Semantic knowledge about regularities in the world exerts a powerful impact on the speed of searches for embedded objects, when compared to searches for isolated objects (for a review see Wolfe et al., 2011). Therefore, it is reasonable to hypothesize that the semantic regularities, provided by a real scene, might also be beneficial in forming memories for targets found in those scenes. In Experiment 3, we examined the role of scene semantics for object memories. Participants searched for or explicitly memorized objects that were superimposed on distinct, but uniform, meaningless backgrounds. Crucially, the objects were neither related to the scene background, nor spatially arranged in a meaningful way.

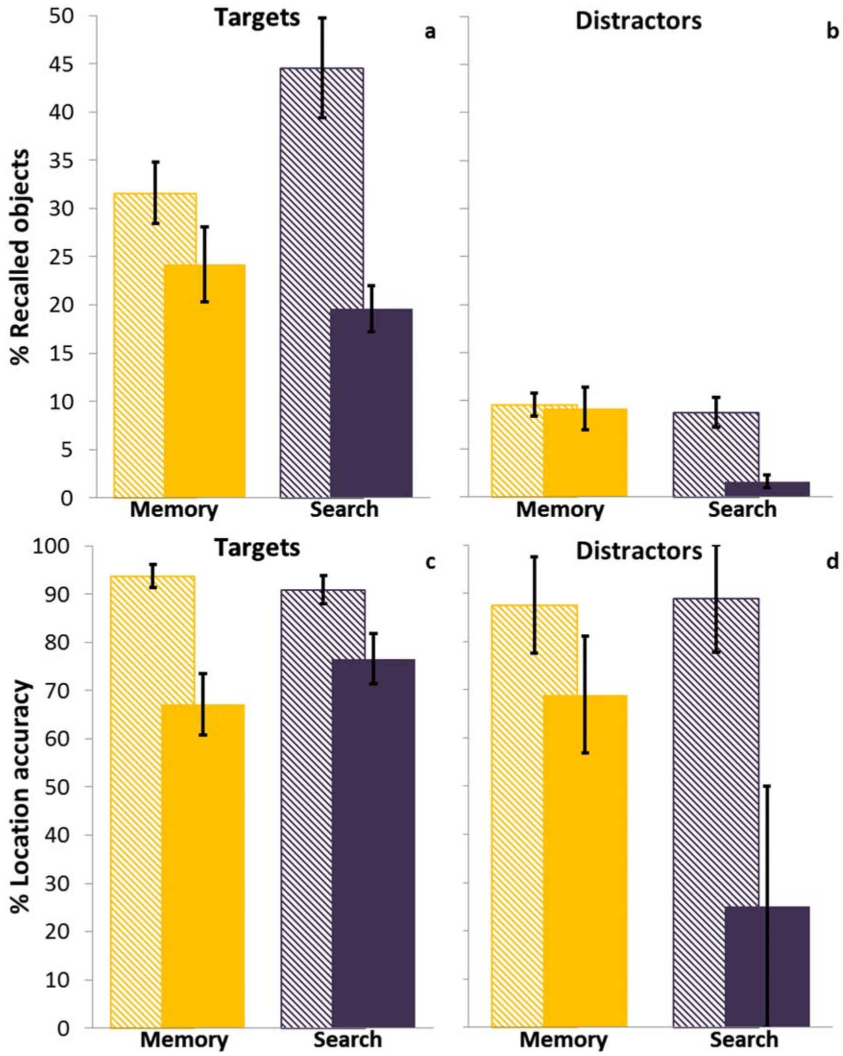

Figure 7. Memory performance as assessed by the mean percentage of recalled objects (a and b) and location accuracy for these recalled objects (c and d) depicted separately for targets (a and c) and distractors (b and d) as a function of Task (Memory=yellow vs. Search= purple). The results of Experiment 2 (objects embedded in scenes) are presented in the according colors, but as dashed lines.

Procedure

The procedure was identical to the procedure of Experiment 2. Only the stimulus material differed in that isolated objects instead of naturalistic scenes were used, as described in the General methods section (see Figure 2e and Figure 2f). Since boundary extension has been found to occur for objects presented on backgrounds lacking scene context (Gottesman & Intraub, 2002; Intraub, Gottesman, & Bills, 1998), we again tested whether participants falsely remembered scene boundaries.

Results

Objects recalled

Contrary to Experiments 1 and 2, there was no recall difference for targets between the Search and Memory condition, $t(9) = 1.09$, $p > 0.3$ (Figure 7a). However, distractors were drawn more frequently after memorization, $t(9) = 3.48$, $p < 0.01$ (Figure 7b). More targets than distractors were reproduced both in the Search, $t(9) = 8.90$, $p < 0.01$ and Memory condition, $t(9) = 4.22$, $p < 0.01$. This time, there was no difference of the magnitude of the effect between conditions, $t(9) < 1$ (Search: M = 18%, SEM = 2% versus Memory: M = 15%, SEM = 3.6%).

Location accuracy

There was a benefit in location accuracy for searched compared to memorized targets, $t(9) = 2.40$, $p < 0.05$ (Figure 7c). Between distractor comparisons were not conducted, as only three participants recalled distractors in all conditions, though the location accuracy

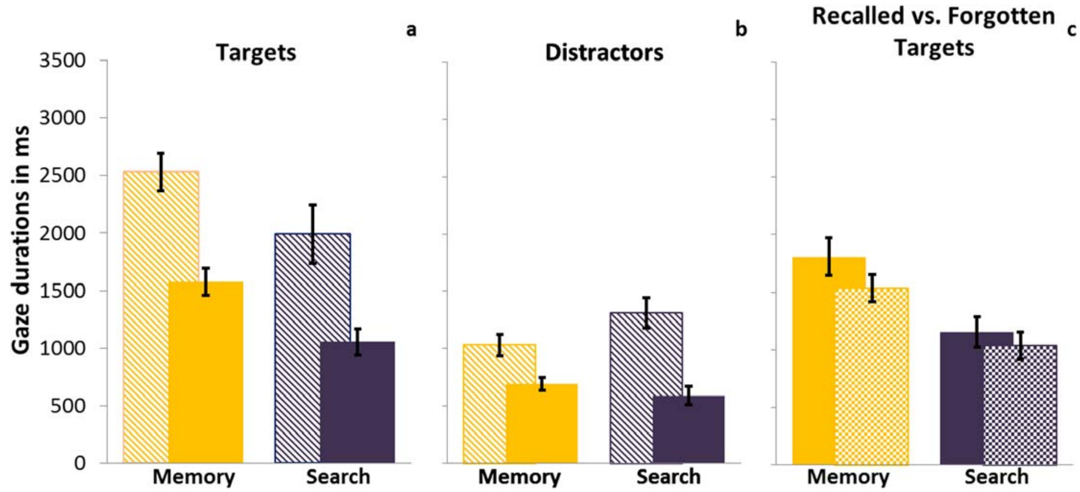

Figure 8. Gaze durations as a function of Task (Memory=yellow vs. Search= purple) for target objects (a) and distractors (b). The results of Experiment 2 (objects embedded in scenes) are presented in the according colors, but as dashed lines. The right graph (c) depicts gaze durations on targets as a function of Task (Memory=yellow vs. Search= purple) and Targets Recalled (Present=full vs. Absent=checkered).

appears to be lower for distractors in the Search condition (Figure 7d).

Boundary extension

Data from one participant was not collected, because of technical issues with the testing machine. In contrast to Experiments 1 and 2, we found reliable boundary extension for Search, $t(8) = 2.48$, $p < 0.05$ (mean boundary extension value = 2.1), but not for Memory, $t(8) = 1.67$, $p > 0.1$ (mean boundary extension value = 2.3). Mean boundary extension values did not differ between conditions, $t(8) = 1.47$, $p > 0.1$.

Task order effects on recall performance

There was no main effect of Task order on the recall performance, $F(8, 1) = 1.09$, $p > 0.05$ and no interaction between Task and Task order, $F(8, 1) < 1$. Only the main effect of Task became significant, $F(8, 1) = 5.34$, $p < 0.05$. The order of the two tasks neither affected recall in the Search condition, $t(6.47) = 1.22$, $p > 0.05$, nor in the Memory condition, $t(7.93) < 1$.

Gaze durations

Figure 8 depicts the mean gaze durations on objects during search and memorization. Gaze durations were significantly longer for targets during memorization, $t(9) = 3.13$, $p < 0.02$, but did not differ for distractors between tasks, $t(9) < 1$ (Figure 8a and 8b). Participants gazed longer on targets than distractors in both Search, $t(9) = 5.87$, $p < 0.01$ and Memory conditions, $t(9) = 5.60$, $p < 0.01$.

Dwell times differed significantly depending on whether a target was recalled or not, $F(9, 1) = 5.34$, $p < 0.05$, showing that recalled targets received higher gaze durations than targets that were not drawn (Figure 8c). There also was a main effect of Task, $F(9, 1) = 14.02$, $p < 0.01$ on dwell times, but no significant interaction, $F(9, 1) = 1.04$, $p > 0.3$. Because seven participants did not recall any distractor objects in some of the conditions, analyses for drawn distractor were not conducted for this experiment.

Naturalistic scenes (Experiment 2) vs. isolated objects (Experiment 3)

To directly test whether the lack of search superiority in Experiment 3 was due to an increase in performance for memorized objects or a decrease of performance for the ones searched in scenes that lack scene semantics, we employed between-subject comparisons (Experiment 2: Present vs. Experiment 3: Absent) for both Memory and Search conditions.

Objects recalled

Comparing the Search conditions of Experiments 2 (where scene semantics were present) and 3 (scene semantics absent), we found a substantial decrease of performance when scene semantics were not available to guide search, $t(12.6) = 4.40$, $p < 0.01$ (Figure 7a). However, scene semantics had no significant impact on memory for targets in the Memory condition, as there was no reliable difference between the two experiments, $t(17.4) = 1.48$, $p > 0.1$.

Similar to the results for targets, distractor recall after search was reduced when comparing Experiment 2 (scene) with Experiment 3 (isolated objects), $t(12.1) = 4.27$, $p < 0.01$. There was no such difference between experiments for the Memory condition, $t(13.9) < 1$ (Figure 7b).

Location accuracy

For location accuracy we found that more of the recalled target objects were placed accurately after search in scenes (Experiment 2) than among isolated objects (Experiment 3), $t(14) = 2.37$, $p < 0.05$ (Figure 7c). There was also a significant difference in location accuracy between the two experiments after memorization, with a benefit for Experiment 2 $t(11.44) = 3.91$, $p < 0.01$.

Gaze durations

Figure 8 depicts the mean gaze durations on objects during search and memorization for Experiment 2 and 3.

Gaze durations on targets during Experiment 3 were overall reduced compared to Experiment 2 both for the Search and the Memory condition ($t(12.56) = 3.91$, $p < 0.01$ and $t(16.27) = 4.70$, $p < 0.01$, respectively). Crucially, the reduction in gaze durations was almost identical in both conditions (–935 ms for the Search and –954 ms for the Memory condition).

Similarly, gaze durations on distractor objects were reduced without scene semantics for the Search, $t(14.66) = 4.63$, $p < 0.01$ and Memory condition, $t(14.69) = 3.11$, $p < 0.01$.

Discussion

In Experiments 1 and 2, we demonstrated significant recall benefits for objects in the search task in comparison to the explicit memorization task. In Experiment 3, where participants searched for and memorized isolated objects superimposed on uniform backgrounds that lack scene semantics, the recall benefit for searched objects disappeared (Figure 7). We see a large drop for recall in the Search condition of Experiment 3 compared to Experiment 2 (–25%). This is much more pronounced than the drop for the Memory condition (–8%).

This result suggests that scene semantics not only benefit search by efficiently guiding attention, but also support formation of stronger memory representations for the searched objects in scenes (Experiment 2) compared to objects that were searched for without the involvement of scene semantic guidance (Experiment 3). The gaze duration difference between Memory and Search remained unchanged between these latter two experiments, minimizing the possibility that dwell time differences play a crucial role in the performance drop for searched objects after the elimination of scene semantics (Figure 8).

General discussion

The aim of the current study was to directly compare the memory representations incidentally formed during search to those intentionally formed during explicit memorization. One could imagine that a participant, conducting a search without the intention to store object information in memory, might be left with a relatively weak representation similar to those produced by free viewing. Accordingly, Tatler and Tatler (2013) have shown that memory performance is weaker when participants are uninformed about a subsequent memory test compared to when they are explicitly told to prepare for that memory test. At the same time, prior search for an object embedded in a real scene, substantially speeds subsequent search for that object (Hollingworth, 2012; Võ & Wolfe, 2012, 2013b), implying strong memory representations for searched objects. The present study now directly tested recall memory–here the ability to draw objects in scenes from memory—after search against explicit memorization. The results show that recall memory is superior for incidentally encoded during search compared to explicitly memorized objects. But what causes this counterintuitive recall superiority for objects of search?

First of all, gaze duration should be considered as a relatively uninteresting factor that could influence the number of recalled objects, given that time spent fixating an item predicts memory performance (Hollingworth & Henderson, 2002). Thus, the enhanced recall performance in the Search condition of Experiment 1 could be attributable (at least partially) to the somewhat longer gaze durations on objects during search. However, after equating dwell times between tasks in Experiment 2, the recall benefit for searched objects remained significant. Thus, although the time spent on an object may influence the overall memory performance, it is differences in the task that seem to produce differences in recall memory. Curiously, recalled target objects embedded in a scene, were not gazed at longer than objects that were subsequently not recalled (Figures 6c). Memory formation seems sensitive to task requirements, rather than the mere time spent on an object. This conclusion falls in line with previous findings that recall of the identity of an object is not predicted by the number of fixations on that object (Tatler et al., 2005; Tatler & Tatler, 2013). Interestingly, when the semantic coherence of the scene was removed, dwell times did predict if a target was recalled or not (Figure 8c). One account for that result could be that the lack of scene semantics led our participants to focus more on the objects themselves rather than acquiring — potentially beneficial — semantic information from the visual periphery and therefore foveal processing as seen in dwell times could have gained importance.

On the basis of Experiments 1 and 2, it could be proposed that the act of searching, itself, generates better memory representations compared to explicit memorization. Perhaps, finding a target object is experienced as rewarding and thus enhances memory. In the Memory condition, where targets are surrounded with a salient frame, target acquisition may be too easy to be rewarding. This hypothesis is rejected by Experiment 3 with its semantically meaningless scenes of isolated objects superimposed on uniform, textured backgrounds. In this case, mere search did not yield a benefit relative to memorization conditions. Significantly, participants recalled only half as many target objects after search for isolated objects in Experiment 3 compared to semantically consistent objects in coherent scenes in Experiment 2, whereas memorization conditions produced comparable results, therefore showing less reliance on scene semantics. This suggests that the scene semantics play a crucial role in supporting more effective formation of scene memories during search.

Further, the decrease in recall in the Search condition of Experiment 3 might be attributed to the decrease in dwell times relative to Experiment 2 (Figure 8) but there is a similar decrease in the Memorization condition of Experiment 3 relative to Experiment 2 without a comparable reduction in recall. Again, it seems more likely that enhanced recall in the Search condition is dependent on the embedding of targets in a meaningful scene.

Turning to the effect of task relevance, we replicated the previous finding that task-relevant objects are recalled more frequently than task irrelevant objects (e.g., Castelhano & Henderson, 2005; Tatler & Tatler, 2013; Võ et al., 2008; Williams et al., 2005). This effect was stronger in Search than in Memorization conditions of Experiments 1 and 2. Notably, this task relevance effect was not significant in Experiment 3, where the difference between recalled targets and distractors was the same between conditions. This insignificance was mainly accounted for by a weaker effect in the Search condition, which could be attributed to the near to floor recall performance on task irrelevant, distractor objects. In any case, further research is needed to uncover the influence of task instructions and scene semantics on task-relevant and irrelevant object memories.

Within semantically meaningful scenes, participants showed close to ceiling performance regarding the location memory for recalled objects. Participants were equally good in recalling the accurate position of task-relevant and task irrelevant objects in both conditions. Meaning, that if an object was recalled, it was very likely to be recalled in its accurate position, irrespective of its relevance or if it was searched for or memorized. However, task instruction had an impact on location memory for the semantically meaningless scene. Performance was better for searched targets, than memorized ones, but this pattern seemed reversed for irrelevant objects. Overall performance was worse for isolated objects, compared to objects embedded in scenes.

In addition to testing object memory, we were also interested in whether task instructions modulated recognition memory for scenes as a whole. We therefore assessed boundary extension, the phenomenon of remembering parts of a scene that were never visually presented (Intraub & Richardson, 1989). We did not find any signs of boundary extension in our first two experiments, where objects were embedded in semantically rich scenes. Participants remembered the space depicted inside the boundaries of the naturalistic image almost perfectly in all conditions. One may speculate that the high complexity of the scenes in our study in combination with the trial-by-trial memorization and search procedure created representations resistant to boundary extension. This is different from the methods of Intraub and Richardson in which simple scenes were viewed only once. Since boundary extension was merely an additional recall measure in this study, strong conclusions would not be warranted. It would be interesting to examine the task dependence of boundary extension but that would require further experimentation.

In the current study, memory performance was assessed by asking participants to draw from memory. Drawings maybe do not represent participants' entire extant scene representation. It might be the case that participants would recognize objects they did not recall in a recognition memory or change detection paradigm setting. However, explicit recall provides a good method for assessing participants' free expression of what they remember. This latitude is advantageous, as no previous assumptions of what participants might remember/forget need to be made in order to construct a recognition paradigm. In consequence, as to not limit the recall of scene representations, we applied an explicit memory test, rather than a restrictive test procedure (for an example of participants drawings see Supplementary Material).

Finally, while we tried to make both the search and the memory tasks as comparable as possible, a complete eradication of differences would not serve the purpose of looking at the effect of task differences on scene memory. Future work might want to address the specific ingredients of incidental encoding during search that make it superior to incidental memorization.

In sum, our findings indicate that searching for objects embedded in real-world naturalistic scenes creates strong and lasting object representations that appear to be stronger than intentionally formed memory representations. We conclude that scene semantics not only help us search for objects in real scenes, but also create scene representations that boost our memory for objects we have sought to find.

Keywords: scene perception, scene search, scene semantics, eye movements, object memory, task dependent representations

Acknowledgments

This work was supported by Office of Naval Research (ONR) Grant N000141010278 to J. M. Wolfe and by National Institute of Health (NIH) Grant F32EY022558 as well as Deutsche Forschungsgemeinschaft (DFG) Grant VO 1683/2-1 to M. L.-H. Võ. We wish to thank Sage Boettcher and Emilie Josephs for their assistance with data collection.

Commercial relationships: none.
Corresponding author: Dejan Draschkow.
Email: dejan.draschkow@yahoo.com; dejan.draschkow@psy.lmu.de.
Address: Ludwig-Maximilians University, Munich, Germany.

References

Boot, W. R., McCarley, J. S., Kramer, A. F., & Peterson, M. S. (2004). Automatic and intentional memory processes in visual search. *Psychonomic Bulletin & Review, 11*(5), 854–61. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/15732694.

Brady, T. F., Konkle, T., Alvarez, G. A., & Oliva, A. (2008). Visual long-term memory has a massive storage capacity for object details. *Proceedings of the National Academy of Sciences, USA, 105*(38), 14325–9. doi:10.1073/pnas.0803390105.

Brockmole, J. R., & Henderson, J. M. (2006). Using real-world scenes as contextual cues for search. *Visual Cognition, 13*(1), 99–108. doi:10.1080/13506280500165188.

Castelhano, M. S., & Henderson, J. M. (2005). *Incidental visual memory for objects in scenes. Visual Cognition: Special Issue on Real-World Scene Perception*, (12), 1017–1040.

Chun, M. M. (2000). Contextual cueing of visual attention. *Trends in Cognitive Sciences, 4*(5), 170–178. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/10782102.

Chun, M. M., & Jiang, Y. (1998). Contextual cueing: Implicit learning and memory of visual context guides spatial attention. *Cognitive Psychology, 36*(1), 28–71. doi:10.1006/cogp.1998.0681.

Gagnier, K. M., & Intraub, H. (2012). When less is more: Line-drawings lead to greater boundary extension than color photographs. *Visual Cognition, 20*(7), 815–824. doi:10.1080/13506285.2012.703705.

Gottesman, C. V., & Intraub, H. (2002). Surface construal and the mental representation of scenes. *Journal of Experimental Psychology. Human Perception and Performance, 28*(3), 589–99. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/12075890.

Hollingworth, A. (2004). Constructing visual representations of natural scenes: The roles of short- and long-term visual memory. *Journal of Experimental Psychology. Human Perception and Performance, 30*(3), 519–37. doi:10.1037/0096-1523.30.3.519.

Hollingworth, A. (2006). Scene and position specificity in visual memory for objects. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 32*(1), 58–69. doi:10.1037/0278-7393.32.1.58.

Hollingworth, A. (2012). Task specificity and the influence of memory on visual search: comment on Võ and Wolfe (2012). *Journal of Experimental Psychology: Human Perception and Performance, 38*(6), 1596–603. doi:10.1037/a0030237.

Hollingworth, A., & Henderson, J. (2002). Accurate visual memory for previously attended objects in natural scenes. *Journal of Experimental Psychology: Human Perception and Performance, 28*(1), 113–136.

Horowitz, T. S., & Wolfe, J. M. (1998). Visual search has no memory. *Nature, 394*(6693), 575–7. doi:10.1038/29068.

Horowitz, T. S., & Wolfe, J. M. (2003). Memory for rejected distractors in visual search? *Visual Cognition, 10*(3), 257–298. doi:10.1080/13506280143000005.

Hout, M. C., & Goldinger, S. D. (2010). Learning in repeated visual search. *Attention, Perception & Psychophysics, 72*(5), 1267–82. doi:10.3758/APP.72.5.1267.

Hout, M. C., & Goldinger, S. D. (2012). Incidental learning speeds visual search by lowering response thresholds, not by improving efficiency: Evidence from eye movements. *Journal of Experimental*

Psychology: Human Perception and Performance, 38(1), 90–112. doi:10.1037/a0023894.

Howard, C. J., Pharaon, R. G., Körner, C., Smith, A. D., & Gilchrist, I. D. (2011). Visual search in the real world: Evidence for the formation of distractor representations. *Perception, 40*(10), 1143–53. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/22308885.

Intraub, H. (2012). Rethinking visual scene perception. *Wiley Interdisciplinary Reviews: Cognitive Science, 3*(1), 117–127. doi:10.1002/wcs.149.

Intraub, H., Gottesman, C. V., & Bills, A. J. (1998). Effects of perceiving and imagining scenes on memory for pictures. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 24*(1), 186–201. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/9438959.

Intraub, H., & Richardson, M. (1989). Wide-angle memories of close-up scenes. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 15*(2), 179–87. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/2522508.

Kit, D., Katz, L., Sullivan, B., Snyder, K., Ballard, D., & Hayhoe, M. (2014). Eye movements, visual search and scene memory, in an immersive virtual environment. *PLoS One, 9*(4), e94362, doi:10.1371/journal.pone.0094362.

Klein, R. M., & MacInnes, W. J. (1999). Inhibition of return is a foraging facilitator in visual search. *Psychological Science, 10*(4), 346–352. doi:10.1111/1467-9280.00166.

Konkle, T., Brady, T. F., Alvarez, G. A., & Oliva, A. (2010). Scene memory is more detailed than you think: The role of categories in visual long-term memory. *Psychological Science, 21*(11), 1551–6. doi:10.1177/0956797610385359.

Körner, C., & Gilchrist, I. D. (2007). Finding a new target in an old display: Evidence for a memory recency effect in visual search. *Psychonomic Bulletin & Review, 14*(5), 846–51. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/18087948.

Körner, C., & Gilchrist, I. D. (2008). Memory processes in multiple-target visual search. *Psychological Research, 72*(1), 99–105. doi:10.1007/s00426-006-0075-1.

Kristjánsson, A. (2000). In search of remembrance: evidence for memory in visual search. *Psychological Science, 11*(4), 328–32. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/11273394.

Kunar, M. A., Flusberg, S., & Wolfe, J. M. (2008). The role of memory and restricted context in repeated visual search. *Perception & Psychophysics, 70*(2), 314–28. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/18372752.

Maxcey-Richard, A. M., & Hollingworth, A. (2013). The strategic retention of task-relevant objects in visual working memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 39*(3), 760–72. doi:10.1037/a0029496.

Peterson, M. S., Beck, M. R., & Vomela, M. (2007). Visual search is guided by prospective and retrospective memory. *Perception & Psychophysics, 69*(1), 123–35. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/17515222.

Peterson, M. S., Kramer, A. F., Wang, R. F., Irwin, D. E., & McCarley, J. S. (2001). Visual search has memory. *Psychological Science, 12*(4), 287–92. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/11476094.

Solman, G. J. F., & Smilek, D. (2010). Item-specific location memory in visual search. *Vision Research, 50*(23), 2430–8. doi:10.1016/j.visres.2010.09.008.

Standing, L. (1973). Learning 10,000 pictures. *The Quarterly Journal of Experimental Psychology, 25*(2), 207–22. doi:10.1080/14640747308400340.

Tatler, B. W., Gilchrist, I. D., & Land, M. F. (2005). Visual memory for objects in natural scenes: From fixations to object files. *The Quarterly Journal of Experimental Psychology. Section A, Human Experimental Psychology, 58*(5), 931–60. doi:10.1080/02724980443000430.

Tatler, B. W., Hayhoe, M. M., Land, M. F., & Ballard, D. H. (2011). Eye guidance in natural vision: reinterpreting salience. *Journal of Vision, 11*(5):5, 1–23, http://www.journalofvision.org/content/11/5/5, doi:10.1167/11.5.5. [PubMed] [Article]

Tatler, B. W., & Melcher, D. (2007). Pictures in mind: initial encoding of object properties varies with the realism of the scene stimulus. *Perception, 36*(12), 1715–29. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/18283923.

Tatler, B. W., & Tatler, S. L. (2013). The influence of instructions on object memory in a real-world setting. *Journal of Vision, 13*(2):5, 1–13, http://www.journalofvision.org/content/13/2/5, doi:10.1167/13.2.5. [PubMed] [Article]

Vickery, T. J., King, L.-W., & Jiang, Y. (2005). Setting up the target template in visual search. *Journal of Vision*, *5*(1):8, 81–92, http://www.journalofvision.org/content/5/1/8, doi:10:1167/5.1.8. [PubMed] [Article]

Võ, M. L.-H., & Henderson, J. M. (2010). The time course of initial scene processing for eye movement guidance in natural scene search. *Journal of Vision*, *10*(3):14, 1–13, http://www.journalofvision.org/

content/10/3/14, doi:10.1167/10.3.14. [PubMed] [Article]

Võ, M. L.-H., Schneider, W., & Matthias, E. (2008). Transsaccadic scene memory revisited: A "theory of visual attention (TVA)" based approach to recognition memory and confidence for objects in naturalistic scenes. *Journal of Eye Movement Research, 2*(2), 13.

Võ, M. L.-H., & Wolfe, J. M. (2012). When does repeated search in scenes involve memory? Looking at versus looking for objects in scenes. *Journal of Experimental Psychology: Human Perception and Performance, 38*(1), 23–41. doi:10.1037/a0024147.

Võ, M. L.-H., & Wolfe, J. M. (2013a). Differential electrophysiological signatures of semantic and syntactic scene processing. *Psychological Science, 24*(9), 1816–23. doi:10.1177/0956797613476955.

Võ, M. L.-H., & Wolfe, J. M. (2013b). The interplay of episodic and semantic memory in guiding repeated search in scenes. *Cognition, 126*(2), 198–212. doi:10.1016/j.cognition.2012.09.017.

Williams, C. C., Henderson, J. M., & Zacks, R. T. (2005). Incidental visual memory for targets and distractors in visual search. *Perception & Psychophysics, 67*(5), 816–27. Retrieved from http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1751468&tool=pmcentrez&rendertype=abstract.

Wolfe, J. M., Alvarez, G. A., Rosenholtz, R., Kuzmova, Y. I., & Sherman, A. M. (2011). Visual search for arbitrary objects in real scenes. *Attention, Perception & Psychophysics, 73*(6), 1650–71. doi:10.3758/s13414-011-0153-3.

Wolfe, J. M., Võ, M. L.-H., Evans, K. K., & Greene, M. R. (2011). Visual search in scenes involves selective and nonselective pathways. *Trends in Cognitive Sciences, 15*(2), 77–84. doi:10.1016/j.tics.2010.12.001.