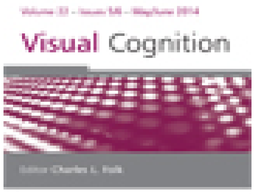

Visual Cognition

ISSN: 1350-6285 (Print) 1464-0716 (Online) Journal homepage: http://www.tandfonline.com/loi/pvis20

Guided search through memory

Maria Nordfang & Jeremy M. Wolfe

To cite this article: Maria Nordfang & Jeremy M. Wolfe (2018): Guided search through memory, Visual Cognition, DOI: 10.1080/13506285.2018.1439851

To link to this article: https://doi.org/10.1080/13506285.2018.1439851

Published online: 06 Mar 2018.

VISUAL COGNITION, 2018
https://doi.org/10.1080/13506285.2018.1439851

Guided search through memory

Maria Nordfang [a] and Jeremy M. Wolfe[b,c]

[a]Department of Neurology, Copenhagen University Hospital - Rigshospitalet, Glostrup, Denmark; [b]Departments of Ophthalmology and Radiology, Harvard Medical School, Boston, USA; [c]Department of Surgery, Brigham and Women's Hospital, Cambridge, USA

ABSTRACT
In "hybrid" search, observers search a visual display for any of several targets held in memory. It is known that the contents of the memory set can guide visual search (e.g., if the memorized targets are all animals, visual attention can be guided away from signs). It is not known if the visual display can guide memory search (e.g., if the memory set is composed of signs and animals, can a visual display of signs restrict memory search to just the signs?). In three hybrid search experiments, participants memorized sets of items that belonged to either one or several categories. Participants were then presented with visual displays containing multiple items, also drawn from one or several categories. Participants were asked to determine if any of the items from their current memory set were present in the visual display. We replicate the finding that visual search can be guided by the contents of memory. We find weaker, novel evidence that memory search can be guided by the contents of the visual display.

ARTICLE HISTORY
Received 9 May 2017
Accepted 5 February 2018

KEYWORDS
Memory search; visual search; visual attention; categories; attentional selection

Visual search has been a significant research topic in the field of cognition for several decades (Wolfe & Horowitz, 2017). Typically, participants search for a single target held in memory in a visual display containing some number of distracting items. Critical variables for the success of such searches include specific features of the target (Wolfe & Horowitz, 2017), target–distractor similarity and heterogeneity of the distractors (Duncan & Humphreys, 1989), and so forth (for reviews, see Carrasco, 2011; Chan & Hayward, 2012; Wolfe, 2015). We know much about the components that influence visual search efficiency when people are looking for a single item. In everyday life, however, search tasks are often more complex than the typical laboratory task. It is not uncommon to search through items in the visual field for any of several possible targets held in memory. For example, one might search the refrigerator for the multiple ingredients needed for tonight's dinner or scan a crowded room for any of many friends, or rummage in the closet for items to pack for an upcoming vacation. Finding a target item in these situations involve searching both the visual field and memory. This combination of visual search and memory search is known as "hybrid search" (Schneider & Shiffrin, 1977; Wolfe, 2012).

Hybrid search tasks have been explored in a limited number of studies. Early work (e.g., Schneider & Shiffrin, 1977) involved small numbers of alphanumeric characters as stimuli. More recently, Wolfe (2012) used search for photorealistic objects which permits a very large increase in the visual and memory set sizes. The 2012 results showed that response times (RTs) increase as a linear function of the visual set size. In contrast, RTs increase as a linear function of the log of the memory set size. The curvilinear effect of memory set size on RTs has been replicated in diverse situations, e.g., with rapid serial visual presentation (Drew & Wolfe, 2014), and with words (Boettcher & Wolfe, 2015) and object categories (Cunningham & Wolfe, 2014) as targets of search.

A variety of models can account for a curvilinear relationship between the number of response categories (or memory set size in hybrid experiments) and RT. Using quite a different memory task, Burrows and Okada (1975) found logarithmic RT functions. In motor tasks with multiple response possibilities, RT rises with the log of the number of response options (Hick's law: Hick, 1952; Schneider & Anderson, 2011). The logarithmic results can arise from any process that eliminates a proportion of total responses on each step in a model. Moreover, the curvilinear/

CONTACT Jeremy M. Wolfe jwolfe@bwh.harvard.edu Department of Surgery, Visual Attention Lab, Brigham & Women's Hospital, 64 Sidney St. Suite. 170, Cambridge, MA 02139-417, USA

logarithmic results can be modelled of the output of a diffusion process (e.g., Gronlund & Ratcliff, 1989; Nosofsky, 1989; Ratcliff & Starns, 2013; Leite & Ratcliff, 2010). The contents of memory in a hybrid search task do not need to be specific images or letters. They can be categories. For instance, we can easily ask if there are any animals, cups, or flags in a display: the memory set size is three, but it is three categories not three specific items (Cunningham & Wolfe, 2014). Moreover, specific items and categories can interact. For instance, an observer could memorize 10 specific animals. In a subsequent visual search, if an attended item is an animal, observers will search through the memory set of 10 animals to determine if this is a target. If it is not an animal, the observer's knowledge that all memory items in the animal category will allow the observer to reject the item on the basis of its categorical status (Cunningham & Wolfe, 2014).

We know that not only the number of items in the memory set, but also the number of categories influence the deployment of attention in visual displays. In this paper, we are particularly interested in the reverse interaction. Is it possible for the categories in the visual display to constrain search through the set of targets held in memory?

Let us consider the four situations illustrated in Figure 1. Suppose that the memory set consists of a set of traffic signs (upper left). If the visual display also consists of traffic signs, then the two sets do not constrain each other. That is, what the observer knows about the memory set does not reduce the number of relevant items in the visual set and vice versa. This is the standard hybrid search condition of, for example, Wolfe (2012) where the memory and visual sets were arbitrary collections of objects. Suppose, however, that the visual display is a mixture of animals and traffic signs (lower right) while the memory set remains all signs. In classic visual search, guidance of visual search by the contents of memory has been thoroughly investigated for the case where there is one item in memory and several items in the visual field (reviewed in Wolfe & Horowitz, 2017). Cunningham and Wolfe (2014) have shown that observers can perform hybrid search for a specific category of targets (e.g., find any animal). In the example given here, categorical information is used differently. We investigate whether the categories in the memory set can guide the visual search. Animals have different basic features than traffic signs (i.e., animals are not simple geometric shapes), thus, if you are looking for any of a set of signs held in memory, attention can be guided away from animals in the visual display and preferentially deployed to the visible signs. This reduces the effective visual set size (from eight to four in the example in Figure 1) and, thus, increases the efficiency of search.

What about the reverse situation? It seems reasonable to ask whether memory search can be guided by the categories of the items present in the visual display. It is clear that memory and visual attention interact. There are studies demonstrating substantial overlap in neural activation for visual search and memory search tasks (Cabeza, Ciaramelli, Olson, & Moscovitch, 2008; Chun & Johnson, 2011; Makino, Yokosawa, Takeda, & Kumada, 2004; see Sestieri, Shulman, & Corbetta, 2017, for a different perspective). Chun and Turk-Browne (2007) have stressed that attention affects episodic and perceptual encoding. They argue that the divide between attention and memory may not be as clear as we may think (see also Craik, Govoni, Naveh-Benjamin, & Anderson, 1996). There is a substantial body of work showing that allocation of visual attention is influenced by both working memory and long-term memory (e.g., Desimone, 1996; Olivers, Meijer, & Theeuwes, 2006; Soto, Hodsoll, Rotshtein, & Humphreys, 2008). Our particular interest is in relationships running in the other direction, namely situations where visual attention and the contents of the external world influence retrieval from memory (see Aly & Turk-Browne, 2017; Chun, Golomb, & Turk-Browne, 2011; Ciaramelli, Grady, & Moscovitch, 2008). Recently, Boettcher, Drew, and Wolfe (2017) found some evidence that it is possible to partition the memory set in hybrid search based on the content of a visual scene. In their work, the background scene told observers that a subset of the items, held in memory, were relevant. In a limited set of conditions, observers seemed able to restrict their memory search to a subset of the memory set.

Returning to Figure 1, suppose that the memory set consists of two animals and two signs (upper right). If the visual display consists of only signs (lower left), can the observer use the visual information to guide memory search to the signs, reducing the effective memory set size? This is a variation on the Boettcher

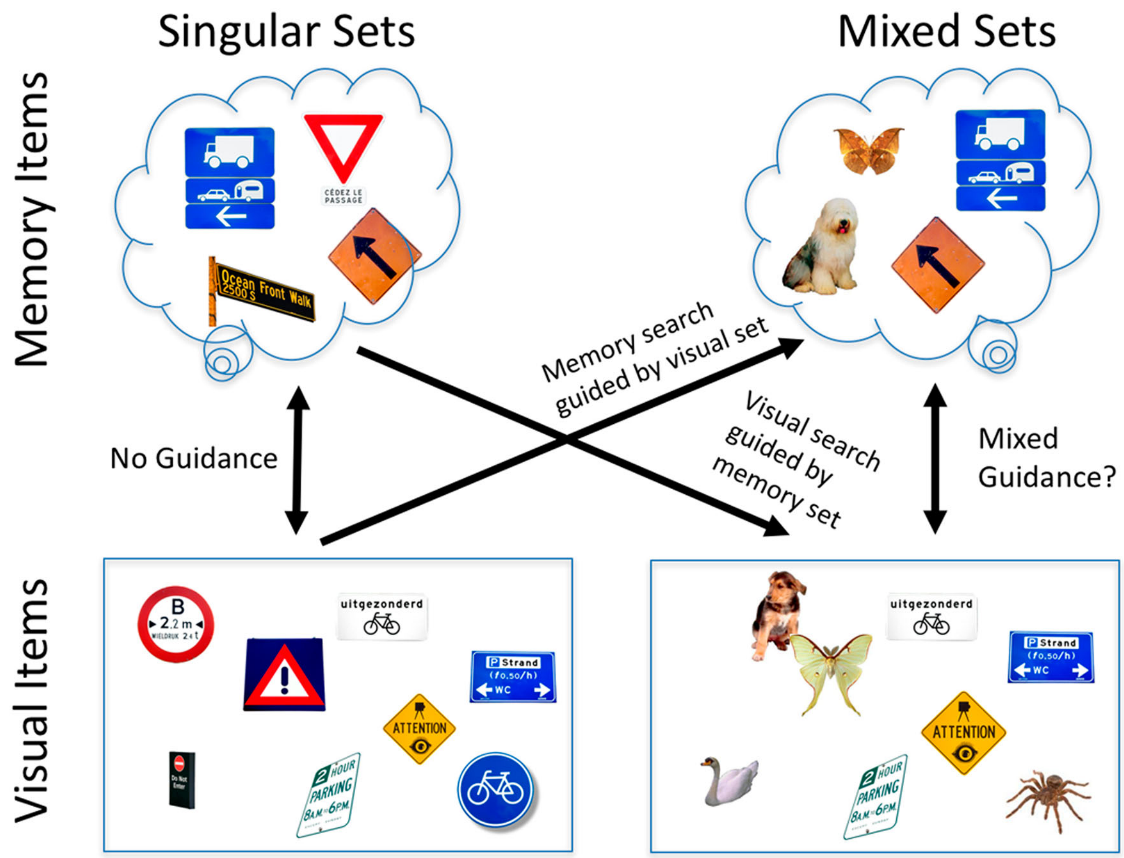

Figure 1. Illustration of the set-up and related hypothesis for Experiments 1 and 2.

et al. (2017) partition experiments. The most novel situation in the current experiments is shown in the upper right and lower right panels of Figure 1. Suppose that both the visual display and the memory set consist of a mix of animals and signs. If the observer attends to an animal in the visual array, after identifying the item as an animal, the observer must decide if it is a target animal. Can the observer restrict memory search to the two animals in the memory set – reducing the effective memory set size – or is the observer forced to search through the irrelevant signs in the memory set as well? A similar effect could run the other direction. If the observer brings an animal into the forefront of memory, can she restrict visual search to animals, reducing the effective visual set size?

The three experiments presented here support the conjecture that memory search can be guided by the categories of items in the visual set (suggestively in Experiments 1 and 2, more definitively in Experiment 3). However, our experiments indicate that this visual guidance of memory search depends on favourable conditions and may not be used by all participants. For visual search, the picture is quite different. Here we see that the categories in the memory set unequivocally guide visual search, replicating earlier studies. Finally, the present experiments fail to find evidence for combined memory and visual guidance. Hybrid search is harder when both the memory set and the visual set are mixed (e.g., animals and signs in both sets), suggesting that a rapid switching between memory subsets is either not possible or not worth the effort from the vantage point of the observer.

Experiment 1

In Experiment 1, we address the question of how visual and memory search interact and whether categorical information from one domain can guide search in the other and vice versa. To investigate this, we presented participants with memory sets comprised of items from either one category or from two conceptually and perceptually distinct categories (see Figure 1 and refer to Table 1 for an example of the stimulus material for one participant. Different participants would see different categories, but the structure

Table 1. Overview of the conditions of Experiment 1, excluding target presence, and an example of the setup for one participant.

Memory Set	Example Memory Set	Visual Categories	Visual Set Size	Example Visual Set
MemSame2	*2 animals*	VisSame	8	*8 animals*
			16	*16 animals*
		VisMix	8	*4 animals and 4 signs*
			16	*8 animals and 8 signs*
MemMix4	*2 plants and 2 cars*	VisSame	8	*8 plants or 8 cars*
			16	*16 plants or 16 cars*
		VisMix	8	*4 plants and 4 cars*
			16	*8 plants and 8 cars*
MemSame4	*4 shoes*	VisSame	8	*8 shoes*
			16	*16 shoes*
		VisMix	8	*4 shoes and 4 fruit*
			16	*8 shoes and 8 fruit*

would be the same). After memorizing these items, participants searched through visual displays with items from either one or both categories. Participants were not explicitly instructed in the categorical nature of the experiment.

We were particularly interested in whether the category of items in a visual display could influence memory search. To investigate this, we looked at hybrid search performance when the memory sets contained mixed-category compared to same-category items and the visual displays contained items from a single category. Referring to Figure 1, consider a visual display containing only signs (lower left). In one memory condition (MemMix4), participants would search through a mixed memory set of two animals and two signs (upper right). This could be compared with search through a memory set with only items from one category (four signs, MemSame4, upper left). If observers can use the categorical status of the visual display to guide search in memory, then only two of the four items in the memory set are relevant in the MemMix4 (upper right) condition. This should be faster than the MemSame4 (upper left) condition where all four signs in memory are relevant. However, if participants were *not* able to use the categorical visual information to guide their memory search then performance should be the same for the trials with mixed and single category memory sets. Furthermore, if this visual guidance of memory was perfect, then performance with a memory set of two animals and two signs (MemMix4) should be as easy as performance with a memory set of just two signs (MemSame2), when the visual display contains only signs.

We were also interested in the reverse comparison: whether a memory set composed of items from one category could be used to guide visual search in a visual set composed of items from two categories (the VisMix16 condition). Performance can be compared to VisSame16 and VisSame8 conditions in which the visual displays are composed of items from a single category. Previous literature suggests that participants can do this kind of visual guidance based on memory contents. Here, these comparisons serve as a useful replication and as a check on the ability of these stimulus materials to enable memory guidance of visual search.

Finally, we wished to see what would happen in the double mixed condition where both memory set and visual display contained two categories (upper right and lower right in Figure 1). However, our predictions are less clear for this condition. On one hand, for these trials, if one attended to an animal in the visual display, in principle, that animal would only need to be compared to the two animals in the mixed memory set. When a sign was attended, only the other two items in memory would be relevant. On the other hand, any such benefit might be lost to the cost of switching between categories in memory search each time a visual item of a new category was attended.

Method

Participants

Twelve paid volunteers (aged 18–47, seven female) took part in the experiment. All participants had at least 20/20 vision with correction and none tested positive for colour blindness on Ishihara's test. All observers gave their informed consent before participating in the experiment.

Stimuli

The stimuli consisted of photorealistic colour images subtending 256×256 pixels used in Cunningham and Wolfe (2014). The images fell into six categories, and each category contained at least 79 exemplars. The categories were animals, signs, cars, plants, shoes, and fruit. The six categories were grouped into pairs that were perceptually and conceptually distinct from each other. Animals were paired with signs, plants with cars, and shoes were paired with fruit. In any specific block of trials, only one of these pairs was used.

Procedure

Participants memorized a set of images and subsequently searched through visual displays for any of the memorized images as in Wolfe (2012). The experiment consisted of three blocks, each using images from one of the three pairs of categories. In each block, participants learned a new memory set. The three memory sets were: two images from the same category (MemSame2), four images from the same category (MemSame4), or two images from each of two categories (four images in total; MemMix4). Across participants, the three different category pairs were used equally often for each type of memory set. For the two MemSame sets, the use of either category from a pair was balanced across participants.

In each block, participants went through a study phase, a memory test, and a hybrid visual and memory search, similar to the procedure in Wolfe (2012). In the study phase, participants were presented with the images from the memory set of that block, one by one, for 3 s each. All images were presented twice. Following the study phase, participants completed a memory test in which all the items of the memory set were probed twice, intermixed with twice as many distractor images of exemplars from the same category. Participants were asked to respond whether the current item belonged to the memory set. Responses were given by key press; the L key for yes, the S key for no. In order to proceed to the hybrid search task, participants had to get at least 95% of the memory test trials correct. If participants made more than 5% errors, all items of the memory set were presented again in a single display that participants could study for as long as desired. The memory test started over from the top, when participants were ready. Two participants went through one of the three memory tests twice, one participant went through one memory test three times. Across participants, the remaining 33 blocks were completed without the need to repeat the memory test.

In the search phase, participants were presented with arrays of visual items and instructed to respond whether an item from the memory set was or was not present. A single item from the current memory set was present on half the trials. Participants were asked to be as fast and accurate as possible. Responses were given by key press. The visual displays contained either eight or 16 items that could come from either one (VisSame) or two (VisMix) categories. This rather complicated design is shown in Table 1. Each participant completed 30 trials per condition equalling 720 trials in total (3 Memory Sets $\times$ 2 Visual Categories $\times$ 2 Visual Set Sizes $\times$ 2 Target Presence $\times$ 30 trials = 720).

Results and discussion

Overall mean RT was 1961 ms. Trials with RTs longer than 6 s or shorter than 200 ms were removed as outliers. This procedure resulted in a loss of fewer than 1% of the trials. Only correct responses (93.4% of trials within the RT cutoffs) were included in the average RT calculations. For all post hoc analysis throughout the paper, p-values are adjusted for multiple comparisons by the method of Hochberg.

Response times

Average RT data are depicted in Figure 2. An omnibus ANOVA with factors Memory Set (MemSame2, MemSame4, MemMix4), Visual Set Type (VisSame, VisMix), Visual Set Size (8, 16), and Target Presence (present, absent) revealed that all main effects and simple interactions, and one three-way interaction (between Memory Set, Visual Set Type, and Target Presence) were significant ($ps < .05$). The results, shown in Figure 2, need to be unpacked in order to evaluate the different types of interactions of visual and memory sets, proposed in Figure 1. Looking first at whether memory search can be guided by a single category in the visual display (e.g., a visual display with all animals) we see that RTs are slightly faster when four items in the memory set are mixed-category (MemMix4, yellow circles) compared to same-category (MemSame4, blue triangles); $t(47) = 2.34$, $p = .024$, $d = 0.34$. However, when MemMix4 is compared with MemSame2, it is clear that the two extra irrelevant items in the mixed memory set place a significant cost on search times; $t(47) = 6.87$, $p < .001$, $d = 0.99$. Thus, when all visual items are from a single category, there is evidence for some ability to prioritize that category in the memory set. However, participants appear not to be able to entirely ignore the other two memory set items.

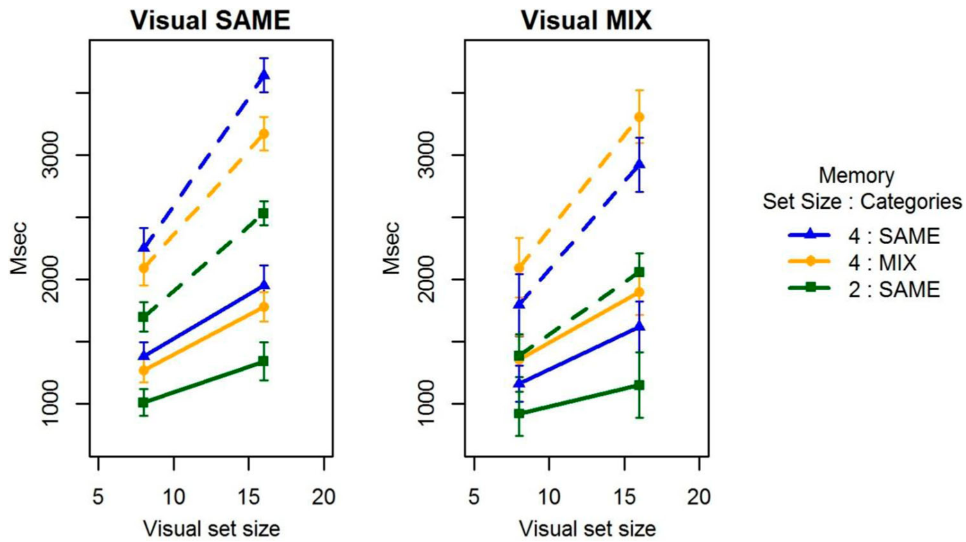

Figure 2. Mean response times for the different memory set conditions and different visual conditions as a function of visual set size in Experiment 1. Broken lines depict target absent trials, full lines depict target present trials. Error bars indicate ± 1 standard error of the mean (within-observer errors calculated by the method of Cousineau, 2005).

When the visual display is a mix of two categories (right side of Figure 2), there is no evidence for an ability to restrict memory search to relevant items in memory. This can be seen by noting that RTs in the mixed memory set (MemMix4, Figure 2 yellow circles) are actually slower than RTs for a memory set of the same size but with just one category (MemSame); $t(47) = -3.40$, $p = .003$, $d = 0.49$. In principle, once the observer selects an item in the mixed visual display, only two items in the mixed memory set would remain relevant. For instance, if the selected visual item was an animal only the two animals in a set of two animals and two signs would be relevant. However, we find no evidence that participants can restrict memory search to those two items.

Our data do show that observers are sensitive to the categorical status of the items in this task. When the visual display is mixed (e.g., animals and signs) and the memory set is drawn from one category (e.g., animals), then observers can guide attention to the visual items that might be members of the target set (attend to animals, not to signs). Thus, if the memory set is a homogeneous set of two or four items from the same category, search is faster when the visual set is divided between two categories. This can be seen by comparing the green square functions between Figures 2(a and b), $t(47) = 7.38$, $p < .001$, $d = 1.06$, or the blue triangle functions, $t(47) = 8.86$, $p < .001$, $d = 1.28$. This is a replication of a similar finding in Cunningham and Wolfe (2014), and tells us that our choice of categories work as expected with regard to guidance of visual search by the contents of memory. We also see that the effect size of guidance of visual search by the contents of memory in this experiment is large or very large (*d*s > 1.00; Sawilowsky, 2009), whereas the effect size of guidance of *memory* search by the contents of the *visual* display is small ($d = 0.34$). However, it is worth noting that, although guidance of visual search by the contents of memory has a large effect on RTs, it too is not perfect. Searching through a visual set of 16 items divided between a relevant and an irrelevant category is slower than searching through eight items that all come from the same category as the memory set, $t(47) = 6.89$, $p < .001$, $d = .99$.

Errors

Error rates were 3–9% with one outlier participant averaging 20% errors. An omnibus ANOVA on errors revealed significant main effects of Visual Set Size, Visual Set Type, and Target Presence. We find significant interactions between Memory Set and Visual Set Type, between Target Presence and all other variables, and between Memory Set, Visual Set Type, and Target Presence (all *p*s < .05, η^2_G ranging from < 0.01 to 0.25). Most of the errors are made when a target is present ($t(12) < .001$, $d = 1.00$), meaning that they are miss errors, if we assume that participants would have found the target if they had searched longer. These miss errors go in the same direction as the

RTs. When the Memory Set contains two categories, participants make fewer errors if the visual display contains just one category. When the Memory Set contains just one category of items, participants make more errors if the visual display also contains just one category. Thus, the harder conditions, with the longer RTs, would have produced even longer RTs if the error rate was reduced. The data show a speed–accuracy covariance, not a speed–accuracy trade-off.

One participant has a high error rate, more than five standard deviations from the mean number of errors for the other participants. In order to determine if errors were altering the pattern of RTs in important ways, we replicated the experiment, using a localization task with a target present on each trial, rather than the present/absent method of Experiment 1. Localization methods tend to minimize false negative/miss errors. In addition, we used larger memory set sizes in an effort to gain more power to see the effects of dividing the memory set between two categories.

Experiment 2

In Experiment 2, we changed the task from a two-alternative forced-choice (present/absent) task to a localization task in which participants used the mouse to click on the location of a target, present on every trial. This reduces the motivation to make speeded guesses. Furthermore, we used larger memory sets of eight or 16 items. All other components of the experiment were similar to Experiment 1.

Method

Participants

Twelve paid volunteers took part in the experiment. Demographic information is missing for two participants; the remaining 10 participants were on average 27.7 years old (range 19 to 55), and five were female, five male. All participants had at least 20/20 vision with correction and none tested positive on Ishihara's test for colour blindness. All observers gave their informed consent before participating in the experiment.

Stimuli and procedure

The stimulus materials and the procedures were the same as in Experiment 1, with the following exceptions. A target from the current memory set was present on every trial. Participants had to click on the target with a computer mouse. Any click within the target image, which was 256×256 pixels, was recorded as a correct response. Each participant completed a total of 960 search trials, yielding 80 observations per condition (3 Memory Sets × 2 Visual Set Sizes × 2 Visual Set Types).

Results and discussion

Trials that were longer than 10 s or faster than 200 ms were counted as outliers. This procedure resulted in a loss of less than 2% of the trials. For the trials within the cutoffs, 2.3% were errors. RT data are depicted in Figure 3.

Again, we started our analysis with an omnibus ANOVA. For this experiment the factors were the three types of Memory Set (MemSame8, MemSame16, MemMix16), two Visual Set Sizes (8, 16), and two Visual Set Types (VisSame or VisMix). The ANOVA revealed that all main effects and simple interactions were significant (all $ps < .05$), and the three-way interaction between Memory Set, Visual Set Type, and Visual Set Size was also significant, $F(2, 22) = 17.06$, $p < .001$, $\eta^2_G = .01$.

As can be seen in Figures 3(a and b), the patterns of results are similar in Experiment 2 to those seen in Experiment 1. Again, we were particularly interested in whether participants could restrict memory search to the relevant subset of the memory set. We see a hint of such an effect for visual displays with items from a single category. In Figure 3(a), the MemMix16 Mixed RTs (yellow circle) appear to be faster than MemSame16 (blue triangle). The effect looks quite promising (175 ms on average). However, it fails to rise to statistical significance $t(23) = 1.84$, $p = .078$, $d = .38$. Nevertheless, comparing the MemMix16 with MemSame8 reveals that these two conditions also are not significantly different from each other, $t(23) = 2.12$, $p = .078$, $d = .43$, whereas MemSame16 was significantly slower than MemSame8, $t(23) = 4.05$, $p = .001$, $d = .83$. Thus, we see a hint of a benefit for the mixed memory sets, but the results are not very convincing.

Turning to Figure 3(b), we see that having one compared to two categories in the Memory Set significantly influenced RTs when the visual displays were

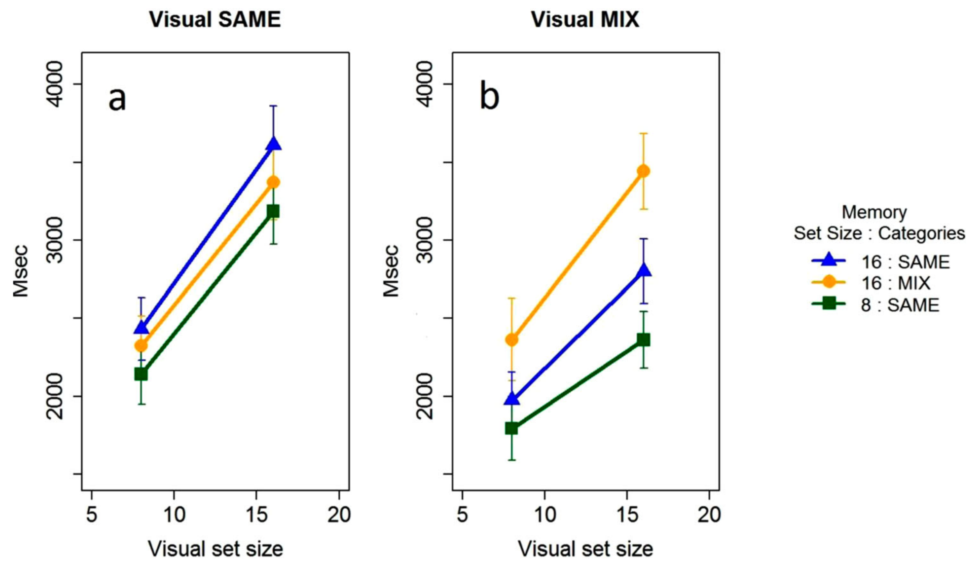

Figure 3. Mean response times for the different memory set conditions and different visual conditions as a function of visual set size for Experiment 2. Error bars indicate ± 1 standard error of the mean (within-observer errors calculated by the method of Cousineau, 2005).

mixed, but, as in Experiment 1, in a negative manner, $t(23) = -6.39$, $p < .001$, $d = 1.30$. As in Experiment 1, RTs were *slower* for the MemMix sets [yellow circles in Figure 3(b)] compared to the MemSame sets (blue diamonds), on average 514 ms slower. Far from improving, search is more difficult when both the visual display and the memory set contain two categories. It might make more sense to think of this as evidence for strong guidance of the visual search by the contents of memory. As in Experiment 1, search through a mixed visual display is faster when the memory set items come from a single category. This is seen by comparing blue diamond or green square functions between Figure 3(a and b). As before, if you know you are looking for specific animals, there is no need to attend to street signs that could never be animals. When this comparison is tested, the number of Visual Categories had a significant effect on RTs when the single category, MemSame sets could guide visual search, $t(47) = 13.79$, $p < .001$, $d = 1.99$. Search efficiency (RT x set size slope) was also significantly influenced by the number of Visual Categories, $t(23) = 7.97$, $p < .001$, $d = 1.63$. Visual search was faster and more efficient when half the items in the visual display did not come from a memory set category.

Overall, both Experiments 1 and 2 replicated the earlier finding that visual attention can be guided to one of two categories in the visual display when the memory set mandates that only one category of items is relevant. The experiments also suggest that it is possible to guide memory search by the contents of the visual display. However, this guidance is relatively weak when averaged over all observers, not rising to statistical significance in Experiment 2.

Can we improve on the weak evidence that memory search can be guided by the categorical status of items in the visual field? In Experiments 1 and 2, the memory set was constant over a whole block of trials, but the number of categories in the visual displays varied from trial to trial. Consequently, the utility of memory guidance changed on a trial-to-trial basis in both experiments. On a given trial, participants did not know in advance whether or not it was advantageous to filter the memory set based on the items in the visual display. It is possible that this setup did not provide sufficient incentive for all participants to initiate categorical memory guidance. We thus decided to run a third experiment with a design that sought to increase observers' motivation to let their memory searches be guided by categorical visual information.

Experiment 3

In Experiment 3 there were no mixed category visual displays. There was always only one category of items in the visual display, though that category could change from trial to trial. Thus, after identifying the categorical status of any one item in the current visual display, participants would know the category of all items in that display. In the MemMix conditions, when there were multiple categories in memory, participants would know which part of their memory set would be relevant for that trial. In this situation, it should always be beneficial to guide memory search by the contents of the visual display. Furthermore, we used a new set of stimuli, deliberately comprised of items from four extremely different categories, in order to encourage participants to use the visual cue to limit memory search to a subset of the memory set. In Experiments 1 and 2, all stimuli, irrespective of category, had been photorealistic images of single objects. For the new Experiment 3, the four categories were drawings, alphanumeric characters, photos of animals isolated on a white background, and photos of scenes. We also increased the number of possible categories in each memory set to a maximum of four different categories. As a result, since each visual display contained only one of the four categories, there could be a very large (4×) benefit to successfully restricting memory search to the relevant memory subset.

Method

Participants

Twelve paid volunteers took part in the experiment. Participants were on average 31.8 years old (range 18 to 55), and six were female. All participants had at least 20/20 vision with correction and none tested positive for colour blindness on Ishihara's test. All gave informed consent.

Stimuli

The stimuli consisted of items from four different categories: keyboard characters, simple drawings, photos of animals on white backgrounds, and photos of scenes. Each category contained 63 unique items. The character stimuli consisted of letters, digits, and symbols drawn in different fonts in Photoshop. The 63 characters were: 123456789ABCDEFGHIJKLMOPQRSTUVWXYZ>[~}@&$=!#*(-%+?";/^÷€∞§£√πΣψ. The simple drawings consisted of custom shapes from Photoshop and icons such as a wineglass, a wrench, and a globe also drawn in Photoshop.[1] The drawings were black, grey, and white. The animal photos were a subset of the stimuli used in the first two experiments. The scene photos were a subset of the stimuli available from Talia Konkle's webpage (see Park, Konkle, & Oliva, 2015).

Procedure

The procedure was the same as in Experiment 1, with the following exceptions. Participants completed a total of nine blocks, each with a different set of four, eight, or 16 items to remember. In a block, these items could either come from one, two or four different categories. All participants went through all nine combinations of three memory set sizes and three possible numbers of memory categories. See Table 2 for an overview of this setup and an example of the stimuli used for one participant. For each participant, the use of the four stimulus categories was approximately balanced. A perfectly balanced use of the categories was not possible in this setup, with four different types of categories and nine different memory set combinations. Across participants, however, each of the four stimulus categories were used equally often for the nine different types of memory sets. Also, any item in a category could only appear once as a memory set item. Memory set items from one block never appeared as visual distractors in another block.

As in the previous experiments, participants had to get at least 95% correct in an old/new recognition task before moving on to the hybrid search task. Four participants reached this criterion for all memory sets after just one presentation of each memory set. Seven participants had to repeat one or two of the nine memory sets in order to reach the criterion. One participant went through three repetitions of

Table 2. Overview of the conditions of Experiment 3, excluding target presence, and an example of the setup for one participant.

Memory Set Size	Memory Categories	Visual Set	Example Memory Set	Example Visual Set
4	1	8 same	4 scenes	8 scenes
	2	8 same	2 drawings and 2 animals	8 drawings or animals
	4	8 same	1 drawing, 1 character, 1 scene, and 1 animal	8 drawings, characters, scenes, or animals
8	1	8 same	8 animals	8 animals
	2	8 same	4 scenes and 4 characters	8 scenes or characters
	4	8 same	2 drawings, 2 characters, 2 scenes, and 2 animals	8 drawings, characters, scenes, or animals
16	1	8 same	16 characters	8 characters
	2	8 same	8 drawings and 8 scenes	8 drawings or scenes
	4	8 same	4 drawings, 4 characters, 4 scenes, and 4 animals	8 drawings, characters, scenes, or animals

one of the memory sets, and one went through four repetitions of a single memory set before reaching the criterion.

On each trial, eight items from the same category were presented in the visual display. On half the trials, one of these items was an item from the current memory set. On the other half, no item from the memory set was present in the visual display. The visually presented items were randomly chosen from one of the categories that was represented in the current memory set. Each participant completed 720 trials, equalling 40 observations per cell (3 Memory Set Sizes × 3 Memory Categories × 2 Target Presence).

Results and discussion

Trials with RTs longer than 8 s or faster than 200 ms were counted as outliers. This corresponded to a loss of less than 15 of all the trials. For the trials within these limits, 9.65 were error trials.

Response times

RT data are depicted in Figure 4. We analysed RTs for the trials for which participants had given a correct answer. An omnibus ANOVA with factors Memory Set Size (four, eight or 16 items in memory), Memory Categories (one, two, or four equal-size categories), and Target Presence (present, absent) revealed that all three main effects were significant ($ps < .001$, η^2_Gs from .08 to .43). The interactions between Memory Categories and Target Presence, and between Memory Set size and Target Presence, were also significant ($ps < .005$, η^2_Gs = .01).

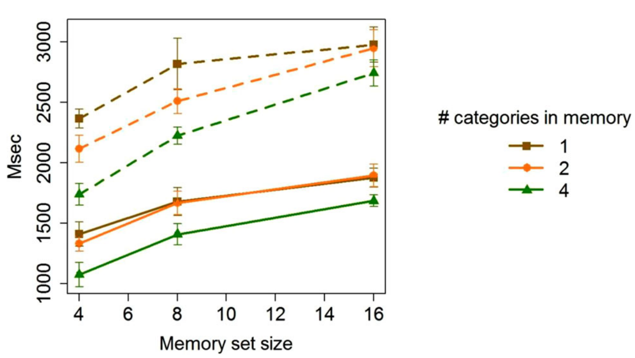

Figure 4. Mean response time data for Experiment 3. The broken lines depict target absent trials, the full lines depict target present trials. Error bars indicate ± 1 standard error of the mean (within-observer errors calculated by the method of Cousineau, 2005).

The important observation is that search is easier when there are four categories in the memory set: for target-present trials $t(35) = 4.84$, $p < .001$, $d = 0.81$, mean difference = 243 ms, and target-absent trials, $t(35) = 3.81$, $p = .002$, $d = 0.64$, mean difference = 289 ms. Holding two categories in memory was not significantly faster than holding one category in memory: for target-present trials: $t(35) = 0.28$, $p = .778$, $d = .05$, and absent trials, $t(35) = 1.51$, $p = .281$, $d = .25$. There were numerical differences of 23 and 196 ms, respectively, in the expected direction.

Experiment 3 shows a clear benefit of using the current visual category to guide memory search. The previous experiment contained similar conditions where visual displays consisted of items that all came from the same category. However, in Experiment 3, all visual displays were single-category displays. This seems to have been the change that persuaded or enabled participants to use the visual information to guide search of the memory set.

Figure 5 provides graphical evidence that this guidance is not perfect, it shows RT as a function of the "effective memory set size". The effective memory set size for each condition is defined as the number items in the actual memory set that would be relevant

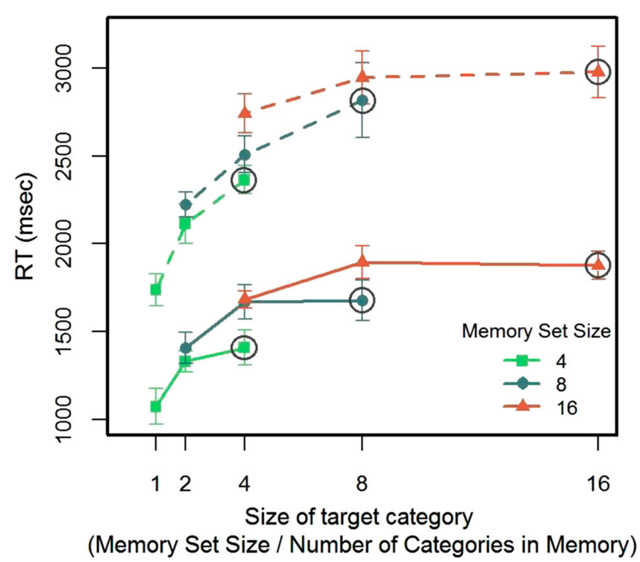

Figure 5. Response time as a function of the effective memory set size in Experiment 3. Effective set size is the number of items in memory that are relevant on the current trial. Broken lines depict data from target absent trials; full lines depict data from target present trials. The circles show the conditions in which memory set size and effective memory set size are the same, i.e., there is only one category of items in memory and no efficiencies can be gained by subdividing the memory set.

on the current trial. Thus, if there are eight items in the memory set and those items fall into four categories, the effective memory set size is $8/4 = 2$. On any trial in that condition, the visual stimulus could guide attention to just two items in the memory set. Circled data points show performance for memory sets consisting of a single category: memory sets that do not need to be sub-divided. If the ability to partition the memory subset was as good as it could be, all the data points would lie on the logarithmic function, defined by those circled points. The size of the full memory set would have no effect on performance. It can be seen that the actual data appear to deviate from this gold standard. An effective memory set size of four out of eight or 16 items, for example, produces longer RTs than a homogeneous memory set of four. This can be seen in the vertical spread of data points for a given effective set size. These deviations are suggestive but are not statistically reliable. On the other hand, if the participants had been completely unable to use the visual displays to guide their memory search, the functions depicted in Figure 5 would have been perfectly flat. Clearly, this is not the case. It appears that the irrelevant items in memory take up some capacity and thus are not entirely cost free. This is also the case for the more robust visual guidance by the contents of memory. Looking at Figure 3 for example, it can be seen that searching for an animal in a display with eight animals is faster than searching a display with eight animals and eight traffic signs, even though only the eight animals are relevant in either case. In the same way, while there is clearly a benefit for the MemMix sets in Experiment 3, the ability to partition the memory set on the basis of visual input does not appear to be perfect.

Errors

We analysed the distribution of the 8.5% error trials that participants had made by an omnibus ANOVA. All three main effects were significant, Memory Set Size, $F(2,22) = 7.08$, $p = .018$, $\eta^2_G = .07$, Target Presence $F(1,11) = 50.91$, $p < .001$, $\eta^2_G = .40$, and Memory Mix $F(2,22) = 4.90$, $p = .023$, $\eta^2_G = .03$. Paired *t*-test revealed that errors increased when Memory Set Size went from four to eight, $t(71) = 3.71$, $p < .001$, mean difference = 3.4%, $d = 0.20$, when a target was present, $t(71) = 14.19$, $p < .001$, mean difference = 11.7%, $d = 1.37$, and decreased when the number of categories present in the memory set went from two to four, $t(71) = -3.55$, $p < .001$, mean difference = −2.8%, $d = 0.42$. Except for Target Presence, these differences went in the same directions as the RT patterns, i.e., more errors were made when RTs were longer. Thus, we found no indication of speed–accuracy trade-offs related to our experimental hypothesis.

General discussion

We can now return to the four types of interaction between visual search and memory search, illustrated in Figure 1. This paper has been particularly focused on the interaction depicted in the diagonal going from bottom left to top right of Figure 1; namely, the ability to guide memory search based on the category of stimuli in the visual display. If the participant knows that all the items in the visual display are signs, can she restrict memory search to the signs held in memory? In Experiments 1 and 2, the results were consistent with this hypothesis, but the effects were highly variable and, in Experiment 2, not statistically reliable. Closer inspection of our data revealed considerable variation across participants. Some participants showed a benefit for this type of memory guidance, others showed none, and a few showed a cost. We did not see any systematic differences in reaction times, or in the stimulus sets that could explain these variations. This individual variation could be the result of a strategic choice (explicit or implicit). Alternatively, it is possible that there are reliable individual differences or that this is simply a noisy measure. The source of the variation could be an interesting topic for future research.

In any case, under the more favourable conditions of Experiment 3, guidance of memory is more robust. In Experiments 1 and 2, the visual sets varied from trial to trial, and memory guidance by visual information was not clearly beneficial for all visual sets. Consequently, the incentive for doing memory guidance varied from trial to trial. In Experiment 3, this was changed. It was always beneficial to guide memory search based on the visual set for the MemMix sets. Furthermore, up to four categories were used in the memory sets, allowing for a possible 4× reduction in the effective memory set size, and thereby increasing the potential benefit of memory guidance. In Experiment 3, we found clear evidence for a benefit when memory sets were mixed (MemMix). There was still individual variation in how consistently we saw evidence of memory guidance, but all participants revealed a benefit for at least some of the MemMix sets and, across participants, this benefit was statistically significant.

Several accounts of memory search could potentially explain why participants overall are faster when the memory sets are mixed. Not all of these explanations evoke strategic guidance. For instance, Brady and colleagues (e.g., Brady, Konkle, & Alvarez, 2011; Konkle, Brady, Alvarez, & Oliva, 2010) have demonstrated that memory capacity is influenced by the similarity between the items in the memory set. Memory performance is better when the items in memory are conceptually distinct. When half the memory set is very dissimilar from the other half, the inter-item competition will be less and recognition of any item from the memory set should be easier. Such an effect could contribute to the differences in memory search efficiency we see in our experiments. However, the effects of the inter-item similarity of the items in memory should not depend on the contents of the visual displays. In particular, one would think that the trials with a single visual category in Experiments 1 and 2 should have produced the same effects as the single visual category trials (all of the trials) in Experiment 3. This did not happen.

That the inter-item similarity in the memory set does not automatically appear to play a strong role for memory search efficiency might be seen as surprising for several accounts of memory search (e.g., Nosofsky, Little, Donkin, & Fific, 2011; Osth & Dennis, 2015). Looking at choice models that envision memory search as a diffusion towards a decision bound (e.g., Leite & Ratcliff, 2010; Ratcliff, Smith, & McKoon, 2015), one could expect that less similar memory set items would diffuse towards the decision bound at a much slower rate than items from the target category, overall resulting in fewer errors for the MemMix sets. In our experiments, the number of errors is too few for a credible analysis of these. However, the error rates roughly follow the same pattern as the reaction times. Conditions that yield slow responses also, overall, result in more errors. According to diffusion theories, reaction times are hard to influence on a trial-to-trial basis, as they are dependent on how the decision bound is set. For Experiments 1 and 2, where the overall similarity between all the items in the visual display and the memory set varies widely from trial to trial, one could hardly expect consistent effects on reaction times, as this would require a trial-to-trial change of the decision bound. As such, our results are roughly in line with this notion of memory search. However, other explanations are possible.

The present experiments were targeted at simply investigating whether any form of memory guidance by visual stimuli was possible at all. We did not set out with any specific hypothesis on how such memory guidance may take place. The present experiments do not distinguish clearly between memory guidance that might be similar to the guidance we know from visual search and effects of categorical and/or perceptual inter-item similarity. These are questions that are open to future research in the area. From the results of these experiments, we argue that it is possible for visually presented categories to influence memory search, when conditions are sufficiently favourable.

It is worth comparing the somewhat equivocal evidence for guidance of memory search by visual input to the clear guidance of visual search by memory contents (the upper-left to lower-right condition of Figure 1). In Experiments 1 and 2, participants consistently demonstrated a benefit for the mixed visual sets when visual search could be guided by the categorical information in the memory sets. When the memory set was drawn from a single category, this could be used to guide the selection of items in visual search when those visual items came from distinct categories. Thus, in Figure 1, if the memory set is a collection of traffic signs and the visual display contains animals and traffic signs, participants appear to be able to guide visual attention to signs and to ignore animals. This could be seen in the data of Experiments 1 and 2 and supports earlier findings of Cunningham and Wolfe (2014), as well as the categorical search studies of Zelinsky and colleagues that demonstrate visual attention can be guided by broad categories rather than specific exemplars (e.g., Maxfield, Stalder, & Zelinsky, 2014; Schmidt & Zelinsky, 2009). This robust visual guidance by the memory categories in Experiments 1 and 2 stands in clear contrast with the more fragile phenomenon of memory guidance by visual stimuli. The visual guidance results serve as a sanity check, showing that our categories work as expected in the visual domain. The different visual and memory guidance results also indicate that the memory guidance and visual guidance may involve different mechanisms (see Konkle et al., 2010, for a similar distinction between visual and memory search).

We are left with the conclusion that in hybrid search, visual search can be guided by the contents of memory – as we knew before – and that memory search can be guided by the contents of the visual scene if the inducements, provided by the task structure, are sufficiently strong and clear. If you imagine a shopping expedition, hybrid search results show that you can hold a multiple item shopping list in your memory (and search through it logarithmically). If your list contains apples, oranges, and grapes; you probably do not need to attend to the bananas (visual guidance by memory). If you are in the bread aisle, you probably do not need to determine if the package of English muffins is an apple, orange, or grape (memory guidance by vision). However, if you are faced with a mix of fruit and bread and you have a memory set containing fruits and breads, it appears that guidance is not possible. Search will be somewhat slower than it would have been under more organized conditions.

Note

1. Many of the drawings were found on the internet at www.shapes4free and http://lukeroberts.deviantart.com/art/.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This work was supported by Army Research Office [Grant Number R00000000000588]; Foundation for the National Institutes of Health [Grant Number EY017001]; The Danish Council for Independent Research [Grant Number 0602-02587B].

ORCID

Maria Nordfang http://orcid.org/0000-0002-2183-0447

References

Aly, M., & Turk-Browne, M. B. (2017). *Hippocampal memory shapes, and is shaped by, attention*. In D. E. Hannula & M. C. Duff (Eds.), *The hippocampus from cells to systems* (pp. 369–403). Springer International Publishing.

Boettcher, S. E., Drew, T., & Wolfe, J. M. (2017). Lost in the supermarket: Quantifying the cost of partitioning memory sets in hybrid search. *Memory & Cognition*, *46*(1), 43–57.

Boettcher, S. E., & Wolfe, J. M. (2015). Searching for the right word: Hybrid visual and memory search for words. *Attention, Perception, & Psychophysics*, *77*(4), 1132–1142.

Brady, T. F., Konkle, T., & Alvarez, G. A. (2011). A review of visual memory capacity: Beyond individual items and toward structured representations. *Journal of Vision*, *11*(5), 4. doi:10.1167/11.5.4

Burrows, D., & Okada, R. (1975). Memory Retrieval from Long and Short Lists. *Science*, *188*(4192), 1031–1033.

Cabeza, R., Ciaramelli, E., Olson, I. R., & Moscovitch, M. (2008). The parietal cortex and episodic memory: An attentional account. *Nature Reviews Neuroscience*, *9*(8), 613–625.

Carrasco, M. (2011). Visual attention: The past 25 years. *Vision Research*, *51*(13), 1484–1525.

Chan, L. K. H., & Hayward, W. G. (2012). Dimension-specific signal modulation in visual search: Evidence from inter-stimulus surround suppression. *Journal of Vision*, *12*(4), 10–10.

Chun, M. M., Golomb, J. D., & Turk-Browne, N. B. (2011). A taxonomy of external and internal attention. *Annual Review of Psychology*, *62*, 73–101.

Chun, M. M., & Johnson, M. K. (2011). Memory: Enduring traces of perceptual and reflective attention. *Neuron*, *72*(4), 520–535.

Chun, M. M., & Turk-Browne, N. B. (2007). Interactions between attention and memory. *Current Opinion in Neurobiology*, *17*(2), 177–184.

Ciaramelli, E., Grady, C. L., & Moscovitch, M. (2008). Top-down and bottom-up attention to memory: A hypothesis (AtoM) on the role of the posterior parietal cortex in memory retrieval. *Neuropsychologia*, *46*(7), 1828–1851.

Cousineau, D. (2005). Confidence intervals in within-subject designs: A simpler solution to Loftus and Masson's method. *Tutorials in Quantitative Methods for Psychology*, *1*(1), 42–45.

Craik, F., Govoni, R., Naveh-Benjamin, M., & Anderson, N. (1996). The effects of divided attention on encoding and retrieval processes in human memory. *Journal of Experimental Psychology: General*, *125*(2), 159–180.

Cunningham, C. A., & Wolfe, J. M. (2014). The role of object categories in hybrid visual and memory search. *Journal of Experimental Psychology: General*, *143*(4), 1585–1599.

Desimone, R. (1996). Neural mechanisms for visual memory and their role in attention. *Proceedings of the National Academy of Sciences*, *93*(24), 13494–13499.

Drew, T., & Wolfe, J. M. (2014). Hybrid search in the temporal domain: Evidence for rapid, serial logarithmic search through memory. *Attention, Perception, & Psychophysics*, *76*(2), 296–303.

Duncan, J., & Humphreys, G. W. (1989). Visual search and stimulus similarity. *Psychological Review*, *96*(3), 433–458.

Gronlund, S. D., & Ratcliff, R. (1989). Time course of item and associative information: Implications for global memory models. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *15*(5), 846–858.

Hick, W. E. (1952). On the rate of gain of information. *Quarterly Journal of Experimental Psychology*, *4*, 11–26.

Konkle, T., Brady, T. F., Alvarez, G. A., & Oliva, A. (2010). Conceptual distinctiveness supports detailed visual long-term memory for real-world objects. *Journal of Experimental Psychology: General*, *139*(3), 558–578.

Leite, F. P., & Ratcliff, R. (2010). Modeling reaction time and accuracy of multiple-alternative decisions. *Attention, Perception, & Psychophysics*, *72*(1), 246–273.

Makino, Y., Yokosawa, K., Takeda, Y., & Kumada, T. (2004). Visual search and memory search engage extensive overlapping cerebral cortices: An fMRI study. *Neuroimage*, *23*(2), 525–533.

Maxfield, J. T., Stalder, W. D., & Zelinsky, G. J. (2014). Effects of target typicality on categorical search. *Journal of Vision*, *14*(12), 1–1. doi:10.1167/14.12.1

Nosofsky, R. M. (1989). Further tests of an exemplar-similarity approach to relating identification and categorization. *Perception & Psychophysics*, *45*(4), 279–290.

Nosofsky, R. M., Little, D. R., Donkin, C., & Fific, M. (2011). Short-term memory scanning viewed as exemplar-based categorization. *Psychological Review*, *118*(2), 280–315.

Olivers, C. N., Meijer, F., & Theeuwes, J. (2006). Feature-based memory-driven attentional capture: Visual working memory content affects visual attention. *Journal of Experimental Psychology: Human Perception and Performance*, *32*(5), 1243–1265.

Osth, A. F., & Dennis, S. (2015). Sources of interference in item and associative recognition memory. *Psychological Review*, *122*(2), 260–311.

Park, S., Konkle, T., & Oliva, A. (2015). Parametric coding of the size and clutter of natural scenes in the human brain. *Cerebral Cortex*, *25*(7), 1792–1805.

Ratcliff, R., Smith, P. L., & McKoon, G. (2015). Modeling regularities in response time and accuracy data with the diffusion model. *Current Directions in Psychological Science*, *24*(6), 458–470.

Ratcliff, R., & Starns, J. J. (2013). Modeling confidence judgments, response times, and multiple choices in decision making: Recognition memory and motion discrimination. *Psychological Review*, *120*(3), 697–719.

Sawilowsky, S. S. (2009). New effect size rules of thumb. *Journal of Modern Applied Statistical Methods*, *8*(2), 597–599.

Schmidt, J., & Zelinsky, G. J. (2009). Short article: Search guidance is proportional to the categorical specificity of a target cue. *The Quarterly Journal of Experimental Psychology*, *62*(10), 1904–1914.

Schneider, D. W., & Anderson, J. R. (2011). A memory-based model of Hick's law. *Cognitive Psychology*, *62*(3), 193–222.

Schneider, W., & Shiffrin, R. M. (1977). Controlled and automatic human information processing: I. Detection, search, and attention. *Psychological Review*, *84*, 1–66.

Sestieri, C., Shulman, G. L., & Corbetta, M. (2017). The contribution of the human posterior parietal cortex to episodic memory. *Nature Reviews Neuroscience*, *18*(3), 183–192.

Soto, D., Hodsoll, J., Rotshtein, P., & Humphreys, G. W. (2008). Automatic guidance of attention from working memory. *Trends in Cognitive Sciences*, *12*(9), 342–348.

Wolfe, J. M. (2012). Saved by a log how do humans perform hybrid visual and memory search? *Psychological Science*, *23*(7), 698–703.

Wolfe, J. M. (2015). Visual search. In J. Fawcett, E. Risko, & A. Kingstone (Eds.), *The handbook of attention* (pp. 27–56). Cambridge, MA: MIT Press.

Wolfe, J. M., & Horowitz, T. S. (2017). Five factors that guide attention in visual search. *Nature Human Behaviour*, 1, 0058. doi:10.1038/s41562-017-0058