



Published in final edited form as:

J Exp Psychol Gen. 2014 August ; 143(4): 1585–1599. doi:10.1037/a0036313.

The role of object categories in hybrid visual and memory search

Corbin A. Cunningham¹ and Jeremy M. Wolfe^{2,3}

¹Johns Hopkins University

²Harvard Medical School

³Brigham and Women's Hospital

Abstract

In hybrid search, observers (Os) search for any of several possible targets in a visual display containing distracting items and, perhaps, a target. Wolfe (2012) found that responses times (RT) in such tasks increased linearly with increases in the number of items in the display. However, RT increased linearly with the *log* of the number of items in the memory set. In earlier work, all items in the memory set were unique instances (e.g. this apple in this pose). Typical real world tasks involve more broadly defined sets of stimuli (e.g. any “apple” or, perhaps, “fruit”). The present experiments show how sets or categories of targets are handled in joint visual and memory search. In Experiment 1, searching for a digit among letters was not like searching for targets from a 10-item memory set, though searching for targets from an N-item memory set of arbitrary alphanumeric characters was like searching for targets from an N-item memory set of arbitrary objects. In Experiment 2, Os searched for any instance of N sets or categories held in memory. This hybrid search was harder than search for specific objects. However, memory search remained logarithmic. Experiment 3 illustrates the interaction of visual guidance and memory search when a subset of visual stimuli are drawn from a target category. Furthermore, we outline a conceptual model, supported by our results, defining the core components that would be necessary to support such categorical hybrid searches.

Corresponding author contact information: cunningham@jhu.edu.

Contact information for Authors:

- Corbin A. Cunningham
Email: cunningham@jhu.edu
Affiliation: Johns Hopkins University
Mailing Address: *Attention and Perception Lab*, Johns Hopkins University, 144 Ames Hall, 3400 N Charles St, Baltimore MD 21218
Phone: 360.271.6306
- Jeremy M. Wolfe
Email: jwolfe@partners.org
Affiliation: Brigham and Women's Hospital / Harvard Medical School
Mailing Address: *Visual Attention Lab*, Brigham & Women's Hospital, 64 Sidney St., Cambridge, MA. 02139
Phone: 617-768-8818
Fax: 617-768-8816

Introduction

In our daily lives, we perform visual search tasks all the time (e.g. Where are my car keys? Can I find the items on my grocery list?). Introspection suggests that tasks like these are reasonably easy. For example, given a mental list of groceries items (e.g. fish, celery, milk, cheese, and eggs), we can successfully search the grocery store. Once we come across any one of those 5 items, we quickly and accurately determine that it was on our list, select it, and continue on to find the rest of the items. But how is this done? How do we keep all of those items in memory and actively search of all of them at once? Prior work (Wolfe, 2012) has addressed this question when the search items are specific instances of specific objects (This cow in this pose.). However, in the grocery example or, indeed, in many real world searches, you are looking for any instance of a target type or category. The grocery list specifies a bunch of celery, not this precise bunch of celery. Moreover, while celery forms a relatively homogeneous set of visual stimuli, “fish” and “cheese” are more visually diverse categories of items, appearing in a wide range of sizes, shapes, and colors. Do the rules uncovered in search for highly specific items apply to the more realistic case of search for any instance of an object or category of objects? Our goal in this paper is to investigate the interactions of visual search and memory search with object recognition and object categorization. To anticipate the results, our core finding is that, as in search for multiple specific objects, the time required to search for multiple types or categories is a linear function of the number of items in the visual display and a linear function of the *log* of the number of items in the memory set. This suggests that, for purposes of hybrid search, a target is a target, whether it is highly specific or a more broadly defined.

Most of the classic research on visual search has involved simple stimuli in simple tasks such as searching for a red item among blue items or searching for instances of the letter, T, among distracting L's. However, in real world search tasks, observers may be searching for multiple targets held in memory. We will call these combinations of visual and memory search “hybrid search” tasks. For example, imagine that you informed your Facebook community about an event. Later, you are at the event and want to determine if any of those Facebook friends (the memory set) are present (the visual search). This hybrid search is search for any of N possible targets (the memory set size) in a visual array containing K items (the visual set size). Previous research has shown that, over a wide range of stimuli, response times (RT) are an essentially linear function of the visual set size (Treisman & Gelade, 1980; Wolfe, 2012). More recently, Wolfe (2012) showed that RTs increased linearly with the *log* of the memory set size. Earlier hybrid search tasks didn't reveal this log function, probably because they used relatively small numbers of alphanumeric items (Cousineau & Larochelle, 2004; Neisser, Novick, & Lazar, 1963; Schneider & Shiffrin, 1977) though (Burrows & Okada, 1975) reported logarithmic RT function in a rather different memory task and Hick's law proposes a log function for choice among multiple actions (Hick, 1952; Schneider & Anderson, 2011).

The Wolfe (2012) experiments took advantage of the massive capacity of picture recognition memory (Brady, Konkle, Alvarez, & Oliva 2008; Konkle, Brady, Alvarez, & Oliva 2010; Shepard, 1967; Standing, Conezio, & Haber, 1970) and had Os searching for 1–

100 specific photographs of objects. More naturalistic hybrid searches are likely to involve less precise specification of the targets. Rather than looking for This Specific Stuffed Bear in This Pose, you are more likely to look for targets defined as “the stuffed bear” or, more broadly, “any stuffed bear”, “any stuffed animal”, “any toy” or some other plausible category of target (Yang & Zelinsky, 2009).

All categories are not created equal. Previous research has shown that processing of items belonging to a specific category depends on the type of category that the items belong to. Basic or entry level categories (e.g. is it a bird or fish?) are distinguished from superordinate level categories (e.g. is it an animal or plant?) and subordinate categories (e.g. is it red-tailed hawk or a chicken?) (Rosch, Mervis, Gray, Johnson, & Boyes-Braem, 1976) and many others since have shown that verification of category membership is faster for basic level categories (e.g. chair, flag, shoe) than for superordinate or subordinate categories. This can be seen as evidence that items are first processed at the basic level with coarser or fine categorization occurring at a later stage – hence the idea of an “entry” level category (Jolicoeur, Gluck, & Kosslyn, 1984). However, it is also possible that the different speeds reflect a single categorization process requiring different accumulations of evidence for different types of categorizations. That is, it may be easy to accumulate the information that differentiates a bird and a house. More subtle features may be needed to determine if the bird is a duck or a goose (Mack & Palmeri, 2011b). Moreover, the categorization of categories is not always obvious and is not immutable. For example, experts can come to treat subordinate categories as entry level (Tanaka & Taylor, 1991). For present purposes, what is important is the distinction between categories – at any level – and specific instances of specific objects. Do the rules of search for those specific instances apply in the more general case? Consequently, in the present work, we use common categories (e.g. fruit, jewelry, animals) without specific attention to their status as basic, superordinate, or subordinate. We find that the specific rules do apply in the general case and we would hypothesize that, while there may be quantitative differences between search for, for example, members of superordinate vs subordinate categories; there will not be qualitative differences between those searches.

Hybrid search can also shed some light on the interaction of visual search with memory systems; notably, in this case, mechanisms of long-term memory. There is a substantial body of work on the interaction of working memory and visual search (Poole & Kane, 2009; Sobel, Gerrie, Poole, & Kane, 2006). Holding an item in working memory influences the course of visual search (Balani, Soto, & Humphreys, 2010; Luria & Vogel, 2011; Olivers, Peters, Houtkamp, & Roelfsema, 2011; Soto, Heinke, Humphreys, & Blanco, 2005) (and vice versa (Eriksen, Eriksen, & Hoffman, 1986)). This leads to the hypothesis that the representation of the target (“search template”) resides in working memory, perhaps in an “active” subset of that working memory (Beck, Hollingworth, & Luck, 2011; Olivers et al., 2011).

The Wolfe (2012) data pose a challenge for this view. Working memory is profoundly capacity-limited (Cowan, 2001). However, that limit does not seem to apply to the number of possible targets in a visual search tasks. The Wolfe (2012) data show that Os have no particular problem searching visual displays for as many as 100 specific targets. No model

of working memory can accommodate 100 templates. Presumably, those 100 items are being held in some sort of long-term memory (LTM). That is certainly the assumption of the research on massive picture memory (Brady et al., 2008). Moreover, the Wolfe (2012) results suggest that the items in memory are sequestered in a separately searchable aspect of LTM. After all, RTs in hybrid search proved to be quite precisely dependent on the number of items in the memory set; not on the overall size of LTM, even if that could be specified. Note that changing from a search for any of 16 items to a search for any of 100 items will change RTs in a predictable manner but will not noticeably change the size of LTM. Moreover, if you are holding 16 items in the memory set for one block of trials, it is possible to switch, with apparently little cost, to 8 different items on the next block, again suggesting an intermediate state of memory that is not classical Working Memory nor is it simple Long Term Memory.

There are a variety of terms that could be used when referring to the type of memory that holds the memory set in a hybrid search task. One could consider this to be a form of episodic LTM (Tulving, 1972; Tulving & Thompson, 1973) in the sense that you are using a memory for *This Specific Set* rather than some more general, semantic knowledge of objects. Baddeley (2000) formalized this as the “episodic buffer” component in his model of working memory. However, episodic memory usually implies a conscious awareness of the episode in time and that does not seem to apply to memory for 100 pictures. When searching through a visual search display, it isn’t the case that you are remembering the period in time that you last remember activating the memory of one or all of the 100 items, held in memory for the present task. This would suggest that the present phenomena is not encapsulated in the episodic buffer, which is “accessed by the central executive through the medium of conscious awareness” (Baddeley, 2000).

“Intermediate-term memory” (McGaugh, 1966; Rosenzweig, Bennett, Colombo, Lee, & Serrano, 1993) is a term that seems to have specialized meaning among those working on the cellular basis of memory, used primarily in discussions of consolidation. Ericsson and Kintsch (1995) introduced the term “Long-term working memory”. It may well cover the sort of memory involved in hybrid search; however, Ericsson and Kintsch were primarily interested in accounting for situations in which expertise seemed to extend the limits of working memory. Experts who memorize vast numbers of digits or perform feats of mental arithmetic are clearly breaking the usual limits on digit span as are chess experts who can reconstruct meaningful game boards with many more than 7 ± 2 pieces on them (Chase & Simon, 1973) and medical doctors who seem to hold many chunks of information in some sort of working memory as they make a diagnosis. Ericsson and Kintsch’s discussion of long term working memory concerns how these experts acquire domain-specific abilities by organizing the coding and retrieval of larger chunks of information. In the case of hybrid search, there is no particular expertise involved. We all seem to be very good at remembering objects and we all seem to search through a memory set of those objects in a similar manner.

Cowan (Cowan, 1988, 1995) offered the somewhat similar concept of “Activated Long Term Memory (ALTM)” to account for cases where the limit on memory is not 7 ± 2 (or 4 ± 2) (Cowan, 2001). The concept captures the idea of an aspect of working memory that is

outside conscious awareness and that involves intimate ties to LTM. We will use that term here, recognizing that we are taking the term without adopting all of its original theoretical purposes. Unlike Cowan, we are not using ALTM to argue against a classic limited-capacity short-term store. Moreover, according to Cowan (personal communication), memory sets of 100 items like those used here and in (Wolfe, 2012), if held over a long block of trials, “have a longer-term flavor that presses the definition of activated LTM”. In spite of these reservations, we think that Cowan’s ALTM captures the idea that some portion of long-term memory can be activated to serve a current task. Thus, rather than adding yet another term to this literature, we will adopt the term, “ALTM”. Note that we are adopting ALTM as a term of convenience and not as a strong commitment to any particular model of long-term memory.

How many searches make up a hybrid search?

Hybrid search is defined as a combination of visual and memory search. The data of Wolfe (2012) suggested an architecture like that shown in Figure 1.

The observer selects an item in the visual display and tests it against the representation of the target in memory. If there are multiple possible targets, the time required to search the memory set will be a logarithmic function of the number of items in the set. If the visual item is the target, search can terminate. If not, the observer must return for another selection or quit the search, declaring the target to be absent. The data to be presented here will require modification of that model. As will be seen, while this may be a fair description of hybrid search in sets of arbitrary objects, it does not account for the broad categorical top-down information that is available when searching for a category rather than a single or set of arbitrary objects. The Figure 1 model is too simple to account for the hybrid search when objects belong to well-defined sets or categories.

Experiment 1a: Do over-learned memory sets obey the log rule?

As noted, in the original hybrid search experiments, the memory sets were arbitrary sets of objects, assembled purely for the purposes of the experiment. Of course, there are other collections of possible target items that form *less* arbitrary sets than just random objects grouped together, i.e. categories. In Experiment 1, we asked if searching for any of 10 digits or 26 letters is like searching through 10 or 26 arbitrary items. Categorizing a character as a number or a letter would be an example of a superordinate-level categorization. However, when faced with a many letters, deciding that a certain letter (among various other letters and across font sizes) is a “P” would be an example of a basic or entry-level task (Wong & Gauthier, 2007). As we will demonstrate, search for letters and numbers is not like search for arbitrary objects.

There is a substantial literature on search for letters among digits and vice versa, much of it devoted to the question of whether letters “pop-out” among digits and/or vice versa (Duncan, 1983; Egeth, Jonides, & Wall, 1972; Hamilton, Mirkin, & Polk, 2006; Krueger, 1984). We are interested in the somewhat different question of whether the different numbers of letters and digits impacts search by changing the effective number of targets.

Method

Ten observers (Os) (Mean age=22.8, Stdev=8.1, 6 females) gave informed consent and were paid \$10/hr to participate in this experiment. All had at least 20/25 vision with correction, all passed the Ishihara Color Test, and all were fluent speakers and readers of English.

Experimental Sessions were carried out on a Macintosh G4 computer running Mac OS 10.5. Experiments were written in Matlab 7.5 (The Mathworks) using the Psychophysics Toolbox (Brainard, 1997; Pelli, 1997), version 3. Stimuli were presented on 20" CRT monitor (Mitsubishi Diamond Pro 91TXM) with resolution set to 1280 × 960 pixels, and an 85 Hz refresh rate. Observers were placed so that their eyes were 57.4 cm from the monitor. At this viewing distance, 1 cm subtends 1° of visual angle (°).

Stimuli were letters, number, or symbols varying in size from 1° to 3° on a white background. All characters were displayed within a 24.5° × 24.5° region of the screen.

In a block of trials, the O's task was either to detect the presence of any letter or of any number. There were four conditions, run in pseudo-random order across Os.

1. Numbers among letters
2. Numbers among letters and symbols
3. Letters among numbers
4. Letters among numbers and symbols.

Letters were uppercase. The set of letters excluded O and I and the set of numbers excluded 1 and 0. Os were informed of this fact. Consequently, one could debate if the memory set size was 8 or 10 for digits, 24 or 26 for letters. As we will see, it does not matter. Symbols used were '!@#\$\$%^&+?'{}[]<>'.

Os were tested on 30 practice trials and 500 experimental trials in each of the four conditions. Targets were present on 50% of trials. There were 100 trials at each of five visual set sizes: 1, 2, 4, 8, or 16 items. On each trial, a target was selected at random from the target set of letters or numbers. Distractors were selected with replacement from the distractor set. Os were asked to respond as quickly and accurately as possible, using the "A" and "L" keys for "absent" and "present" responses, respectively. Breaks were enforced between blocks. Os could also take breaks within blocks as desired. Accuracy feedback was given after each trial.

In standard visual search experiments, the standard deviation of the estimate of the slope of the RT x set size function is roughly 0.5 of the slope estimate itself. This relationship allows us to calculate power of a comparison between two conditions in terms of the ratio of the slopes in those conditions, compared in a two-sided t-test. Effects of interest in visual search experiments tend to be on the order of 2:1 or greater. A 2:1 ratio between slopes requires 8 observers to yield a power of 0.8 at a 0.05 significance level. We run 10 Os (power = 0.94). Testing 10 Os retains a power of > 0.8 for slope ratios down to 1.70. Effects of a less than 1.70:1 ratio are unlikely to be of great theoretical interest.

Results & Discussion

Figure 2 shows the RT x set size functions for correct target present and target absent trials, averaged across all ten Os. RT outlier cutoff was standardized across all experiments at three standard deviations from the mean for each individual subject.

Overall, this removed between 1–2% of trials. Removing RT outliers is something of a dark art. There are always a few RTs that are clearly the result of missed keys, sleeping observers or other problems but there is rarely an obvious separation between bad and legitimately long RTs. The 3SD cutoff is a widely used method, though it undoubtedly removes some “real” RTs. In the present case, this method produces essentially the same results as an arbitrary filter based on looking at the upper tail of the distribution.

In visual search experiments of this sort, the slope of the RT x set size function is considered to be the most useful index of the efficiency of the search as that slope gives an estimate of the cost of adding an item to the visual display. Other processes like the initial stages of visual processing or the time required for the motor response will add to the overall RT (and, thus, to the intercept of the RT x set size function). The costs of attending to each item and deciding about that item will appear in the search slope (Wolfe, 1998). It is clear that these are inefficient searches (Wolfe, 1998) with no evidence for efficient, “parallel” search for letters among numbers or vice versa (Egeth, Jonides, et al., 1972). Additionally, error rates were modest, averaging 6.6% miss errors and 1.9% false alarms. A repeated measures ANOVA on slopes reveals that the conditions differ from each for present trials ($F(3,27)=6.6$, $p=0.004$, generalized eta squared ($g\eta^2$)=0.154) and absent trials ($F(3,27)=7.9$, $p=0.002$, $g\eta^2=0.092$). (Note: We are reporting generalized eta squared because repeated measures ANOVAs have multiple sources of error, a partial eta squared would not account for these sources, rather only the individual error associated with the given effect. The generalized eta squared accounts for all sources of error in the ANOVA, producing a more conservative estimate of effect size.) However, even though the nominal memory set sizes are very different, there is not much difference between search for numbers and search for letters. Bonferroni-corrected paired t-tests show no difference between search for numbers among letters and letters among numbers, ($t(9) = 1.89$, $p > 0.05$). However, Letter search is slightly less efficient than Number search with added symbols for present trials ($t(9) = 3.11$, $p = .012$). Moreover, the non-significant difference is in the wrong direction, with letter search, with its larger memory set size, being slightly more efficient than number search, in the former case. For absent trials, letter search efficiency does not differ from number search though both are inefficient ($t(9) = 2.7$, $p = .026$, Bonferroni Corrected (α) = .0125). It is important to note that a number of factors could influence the efficiency of search for letters among numbers vs. numbers among letter, such as the range of similarity between the numbers and letters in the font used. However, in this experiment, we are simply trying to determine how visual search for members of the categories, numbers and letters, differs from search for a set of N arbitrary objects.

Figure 3 compares the results of Experiment 1 to the results from the original (Wolfe, 2012) hybrid search experiment. The smaller, circle and square (black and white) data points are the slopes of RT x visual set size functions for individual Os in Experiment 1 of the present

paper, plotted at a memory set size of 10 for numbers and 26 for letters. Large, gray circular data points and the accompanying best fit regression line are derived from the data in (Wolfe, 2012). It shows the linear increase in the slope of the RT x visual set size functions as a function of log₂ of the memory set size. Dashed lines around those points show +/- 95% confidence intervals. (Outline squares are data from Experiment 1b and will be discussed later.)

If letters and numbers behaved like arbitrary objects, the Os' data should lie on or parallel to the Wolfe (2012) function. Clearly, that is not the case. All but two data points lie outside the 95% C.I. Based on the results of Experiment 1, therefore, we can safely conclude that Os do not search through the sets of digits or letters in the way that they might search through a set of 10 or 26 arbitrary objects. The most likely account is that, in a search for a letter among numbers or vice versa, the search is based on the categorical status of the item; *is it a letter or a number?* Apparently, the categorization of an item as a letter or number is not directly dependent on the number of elements comprising the category. This is a fairly obvious point, as it otherwise might take much longer to determine that something came from a very large set like "animal".

Note that we do not mean to suggest that it is surprising that categories like "number" or "letter" do not behave like sets of N objects in hybrid search. Indeed, in a task where observers are asked "Is this item a member of the target category", the performance can be better when the target category is larger (Omohundro & Homa, 1981). Such effects can be accounted for by models like Nosofsky's exemplar-based random walk model (EBRW) (Nosofsky, Little, Donkin, & Fific, 2011). In effect, in the Nosofsky et al (2011) formulation, the category with more exemplars better defines a similarity metric that allows members of the category to be discriminated from non-members. For present purposes, what is important is that the pattern of RTs for decisions about category membership can be quite different from decisions about membership in an arbitrary memory set. One could generalize from this to hypothesize that hybrid search for a set of categories would be qualitatively different from search for a set of specific items. Experiments 2 and 3 show that, while there are quantitative differences, the basic qualitative findings are similar: linear search through the visual set and logarithmic search through the memory set.

Returning to Figure 3, one might propose that the faster RTs for these alphanumeric searches reflect much faster identification of characters than of arbitrary objects. We test this possibility in Experiment 1b.

Experiment 1b: Are alphanumeric characters easier to search than objects?

Is search through alphanumeric characters simply more efficient, in general, than search through objects? In prior work, search through arbitrary objects have been found to be fairly inefficient (Vickery, King, & Jiang, 2005; Wolfe, Alvarez, Rosenholtz, Kuzmova, & Sherman, 2011), while search through alphanumeric characters can be quite efficient (Harris, Shaw, & Bates, 1979; Schneider & Shiffrin, 1977). Experiment 1b rejects the hypothesis that hybrid search through alphanumeric characters is markedly more efficient than hybrid search through objects, at least for these alphanumeric stimuli. Looking forward,

Experiment 3 will provide support for the hypothesis that hybrid search for a number among letters or a letter among numbers is speeded because that task is a category identification task.

Methods

Ten observers (Os) (Mean age=22.2, Stdev=6, 7 females) gave informed consent and were paid \$10/hr to participate in this experiment. All had at least 20/25 vision with correction, all passed the Ishihara Color Test, and all were fluent speakers and readers of English.

In this case, the task was to search for any of N, arbitrarily selected, single alphanumeric characters among other arbitrary alphanumeric characters. At the start of each block of trials, Os were shown a memory set of 1, 2, 4 or 8 items. Thus, for a set of 4, Os might be asked to search for “D”, “T”, “4”, and “W”. Items from the memory sets were presented, one at a time, for 3 second each. Os then took a memory test in which they saw characters and had to label them as belonging to the memory set or not. For a memory set of N items, they saw 2N test trials. Of which 50% were in the memory set. Os needed to be 90% correct on the memory test or it repeated until the observer achieved that level of performance. With the memory set securely in memory, Os were then tested on 50 practice and 300 experimental trials, evenly divided between target-present and targetabsent trials and between five visual set sizes: 1, 2, 4, 8, & 16 items. In all other respects, the experiment was similar to Experiment 1a.

Results and Discussion

In the language of Schneider and Shiffrin (1977), this is a “consistent mapping” task in which the target and distractor sets remain fixed over a block of trials. We used the RT outlier cutoff of three standard deviations from the mean for each individual subject.

Overall, this removed between 1–2% of trials. The critical results are shown as the open/tan boxes in Figure 3. These plot the slopes of the RT x visual set size functions for each of the four memory set sizes tested. It is clear that those slopes lie comfortably *above* the slopes for search for arbitrary objects. This demonstrates that search for these arbitrary alphanumeric characters is, in fact, less efficient than search for arbitrary objects and much less efficient than search for members of the category “number” or “letter”. This is not a speed-accuracy tradeoff. The miss error rates are somewhat higher in this experiment (8%) than in Experiment 1a and are very high when the visual set size and memory set size are at their largest (visual set size 16, memory set size 8, miss error rate = 28%). This is a speed-accuracy covariance and correcting for these errors would only make the slopes of the RT x set size functions steeper. Like the errors and the slopes, the mean RTs are substantially slower when Os search for arbitrary alphanumeric characters compared to arbitrary objects.

Experiment 1b rejects the hypothesis that the slopes in Experiment 1a were shallow because all alphanumeric searches are easy. A more plausible account would be that searching for a member of a well-defined category like “letter” is easier than searching for a member of an ill-defined category – in this case, an arbitrary set of items defined as a target category by the experimenter. This is, perhaps, most clearly illustrated by noting that the slope of a

search for any digit in Experiment 1a is 36 msec/item. This is a nominal memory set of 10 items, 8 if we eliminate 0 & 1, as we did. Compare that slope to the 98 msec/item slope for a search for any of 8 arbitrary characters in Experiment 1b. The efficiencies of the searches for numbers among letters in Experiment 1a and for a consistently mapped set of characters in Experiment 1b are low (slopes are high) compared to the classic findings in, for example, Schneider and Shiffrin (1977). It is not entirely clear why this is the case though the older studies typically use small memory and visual set sizes and more trials. We have typically failed to find efficient search for consistently mapped character sets even when this is as simple as a search for “T”s among “L”s (e.g. Palmer, Fencsik, Flusberg, Horowitz, & Wolfe, 2011). The topic deserves more research.

The basic finding of Experiment 1b is that arbitrary alphanumeric targets behave like arbitrary objects. Experiment 1a shows that the pattern of results is different when the set of targets is not arbitrary. Between the extremes of categorically defined targets and arbitrary targets, prior work has shown that there would be effects of specific target choices. For instance, Egeth, Marcus, and Bevan (1972) have shown that visual search for a target set of three numbers is heavily dependent on which particular numbers are memorized. In one block, all observers searched for the number “1” (memory set = 1) in a stream of numbers, 50% prevalence. In another block, they had observers monitor the stream for either the numbers: “1, 4, & 7” or “1, 2, & 3” (memory set = 3). Again, a target appeared with 50% prevalence. The memory set size functions were very different for the two conditions. For observers who received the memory set consisting of the number “1” in one block and the numbers “1, 2, and 3” in another block, the slopes of the memory set size X RT functions averaged 2 msec per item. However, surprisingly when observers received the memory set consisting of the number “1” in one block and the numbers “1, 4, and 7” in another block, the memory search slopes averaged 24 msec per item. The memory set “1, 2, 3” behaves more like a category than the set “1, 4, 7”.

Experiment Two: Searching for multiple categories

Experiment 1 shows that a target category does not behave like a memory set of its components (e.g. 26 letters). The set of all letters can behave as a single “target” (e.g. is this item a letter? If yes then terminate search.) while an arbitrary subset of N letters behaves like N targets in hybrid search (e.g. is this item a letter? If “yes”, then is this item one from my memory set? If “yes”, then terminate search). Thus, searching for any letter allows the observer to stop his or her visual search at the categorical level without having to check specific items in memory. This is similar, but not identical, to the case of searching for a single letter. As noted earlier, many searches are searches for any member of a set. Are there any birds here? Is there some jam in the pantry? Do each of these sets or categories of targets behave like single, specific items in hybrid search? That is, would search for members of any of N sets or categories look like hybrid search for any of N specific objects? Search for items from multiple categories might be the more typical, real world hybrid search. Consider going to the store to look for orange juice, some fish, and a vegetable. This three-item set constitutes a set of three categories, not three specific items. Those categories might be subordinate (the orange juice), basic (fish), or superordinate (vegetable). We will return to the role of category level in the discussion. For present purposes, the important

point is that each target is defined by some infinite set of possible visual stimuli. In Experiment Two, observers search for multiple categories of targets. (Note that portions of the data in Experiment 2 were previously presented at the 2012 Object Perception, Attention, and Memory (OPAM) Conference (Cunningham & Wolfe, 2012)).

Method

Experiment 2 had 10 observers (Mean age=23.7, Stdev=6.83, 4 females) who were asked to memorize 1, 2, 4, or 8 categories. All gave informed consent and were paid \$10/hr to participate in this experiment. Os had at least 20/25 vision with correction. All passed the Ishihara Color Test and were fluent speakers and readers of English.

Twenty categories were used. The level of a category is not always self-evident but ten of these might be considered superordinate categories: Plants, Furniture, Animals, Weapons, Musical Instruments, Fruit, Clothing, Kitchenware, Electronic Appliances, Rocks and Minerals while the other ten might be considered to be basic categories: Picture Frames, Signs, Flags, Cars, Time Pieces, Shoes, Money, Jewelry, Masks, Sweets. These were chosen to be trivially discriminable from each other by naïve observers. Our interest is in the ability to search for instances drawn from multiple categories. While it might be that there would be differences between basic and other categories, we did not separate these in our experiment. As our supermarket example indicates, real-world hybrid search probably involves a mix of category levels, as well. A minimum of 150 photo objects, with an average of ~300 were available in each category. Images were acquired through a commercially available database (Hemera Photo-Objects Vol. I and II). No category was used as a member of the memory set in more than one block of the experiment for any one observer. However, members of previous target categories could be reused as *distractors* in subsequent blocks.

At the start of a block, the categories in the memory set were presented to the observer as words. Os then took a memory test in which they identified pictures of objects as being in or out of the memory set. Os passed the memory test by achieving greater than 90% correct on two successive tests. Images used in the memory test were not used in the subsequent visual search trials. Having passed the memory test, Os searched visual displays where one and only one of the items in the display was a target, drawn from one of the target categories. Distractors were drawn from ALL of the remaining non-target categories. Thus, a memory set size of 4 might consist of Shoes, Fruit, Furniture, and Jewelry. A target would be any shoe, fruit, piece of furniture, or jewelry. A distractor could be any item from the other 16 categories. Displays had visual set sizes of 4, 8, 16, or 32.

In this experiment, targets were present on 100% of trials and Os made a localization response rather than a present/absent response. The two methods tend to produce similar patterns of RTs. Errors tend to be lower in localization tasks. Localizing responses were made by using the mouse to “click” on that target as rapidly as possible. There were 400 trials per memory set size. The memory set sizes (1, 2, 4, and 8 categories) were presented in a randomized order that was counterbalanced across participants.

Results

The first salient fact about search for members of a set of categories is that it is markedly more difficult than search for members of a set of specific objects. This can be seen in Figure 4, where RTs are plotted as a function of the visual set size.

RTs for Experiment 2 are shown with solid lines. Dotted-line data are the equivalent RTs from search for objects in Wolfe (2012) using the same localization method. For example, the efficiency of search for a member of a memory set of 2 categories is the same as the efficiency of search for a member of a memory set of 8 objects.

In Experiment 2 the RT outlier cutoff was again three standard deviations from the mean for each individual subject. Overall, this removed between 1–2% of trials. Most of the very long RTs are errors. Recall that this is a localization experiment. Os must click on a target. When they forget categories, they must still choose something and that guessing process tends to be slow and incorrect. Even with the localization method, error rates are relatively high in this experiment, rising from about 10% at the low visual and memory set sizes to 20% at the high visual and memory set sizes. The main effect of visual set size on error is significant ($F(3,27)=13$, $p=0.003$, $ges=0.13$). The effect of memory set size is not significant, nor is the interaction. One cause of errors is probably proactive interference from target categories on previous blocks that become distractor categories on the current block. In support of this notion, we find that errors rise over the course of the experiment from about 10% to about 15% (using 50-trial bins, $R\text{-sq}=0.30$, $F(1,33)=14$, $p=0.0006$).

In the RT data, as is evident from Figures 4 and 5, the main effects of Visual Set Size, Memory Set Size and their interaction are strongly significant (Visual: $F(3,30)=262.17(!)$, $p<0.0001$, $ges=0.63$, Memory: $F(3,30)=13.33$, $p<0.001$, $ges=0.102$, Interaction $F(9,90)=7.22$, $p=0.003$, $ges=0.03$).

Though hybrid search for categories is more difficult, it appears to be fundamentally similar to hybrid search for objects. In both cases, the effect of visual set size on RT is basically linear while the effect of memory set size appears to be logarithmic. This can be seen in Figure 5, where RT is plotted as a function of the memory set size. Visual inspection shows the RT x memory set size functions to be non-linear. Linear regression on the RT X memory set size functions produces $r\text{-sq}$ values of 0.77 to 0.92. Linear regression of RT x $\log_2(\text{memory set size})$ produces higher $r\text{-sq}$ in all cases; 0.88 to 1.00. As one way to demonstrate that a linear relationship to $\log_2(\text{memory set size})$ is a good fit to the data, we used the data from memory set sizes 1, 2, & 4 to predict the results for memory set size 8. These predictions are shown as the symbols on the right of Fig 5; shaded figures for a linear fit, open figures for the log-linear fit. The errors for linear prediction are 140, 491, 614, & 1326 msec for visual set sizes 4, 8, 16, and 32, respectively. The errors for the log-linear prediction are much smaller: -6, 38, -28, & 170. The relatively large error at the largest set size may reflect a speed-accuracy tradeoff but, over all, there is a convincingly linear relationship between RT and $\log_2(\text{memory set size})$. In this way, a memory set of categories behaves much like a memory set of objects.

Results show that searching instances of categories is markedly slower than searching for specific pictures of objects; slopes for the categories experiment were on average about 39 msec/item slower than the object version in Wolfe (2012) where the slopes were on average 58.5 msec/item (see Figure 4).

Because target categories from previous blocks could serve as distractors on subsequent blocks, interference from previously memorized categories probably contributed to errors. This hypothesis is supported by an increase in error rates over the course of the experiment. Moreover, under rather different circumstances, multiple categories have been shown to interfere with each other (Evans, Horowitz, & Wolfe, 2011). Whatever the cause, the high miss error rate probably produces a speed-accuracy tradeoff that will distort the RT x visual set size functions and the RT x memory set size functions. Nevertheless, despite the relatively high miss errors, RT x visual set size functions remained essentially linear and, more importantly, RT x Memory set size increased with the log of the memory set size.

Discussion

The original, Wolfe (2012) finding of logarithmic memory search could have been a property restricted to retrieval from our massive memory for specific instances of specific objects (Brady et al., 2008; Standing, 1973). The results of Experiment 2 suggest that logarithmic search is a more general characteristic of search through ALTM. Memory performance for the categories used here is not as good as memory for specific objects. This is not surprising. In Wolfe (2012), the observers had seen the specific objects. In the present experiment, observers had not been exposed to the specific instances of the categories tested. Instead, they were relying on their general semantic knowledge of “animals”, “clothing”, etc. Moreover, the ability to commit categories to memory is different than the ability to commit specific images to memory. Our observers can trivially hold 100 objects in ALTM, while error rates were rising fairly sharply by the time we asked observers to search for any of a mere 8 categories (though, of course, each of those categories encompasses an unknown and large number of specific objects). Nevertheless, RTs rise with the log of the number of categories, just as they rose with the log of the number of objects in earlier work. This logarithmic process seems to be a general feature of searches through ALTM. However, as noted above, visual search for multiple categories in memory was significantly harder (or at least took more time) than searching for the same number of objects. Part of the explanation may have to do with the role of the visual attributes of the target items. It is likely to be easier to search for a specific, red apple than for an unspecified member of the “fruit” category. In addition, an item in the visual display might have a feature that distinguishes it from all target objects. It is harder to have a feature that visually distinguishes an item from a set of arbitrary target categories. In this experiment, we selected a set of distinctive categories without regard for their status as superordinate (e.g. animal) or basic (e.g. cars). Had we used a set of more specific basic (e.g. birds) or subordinate categories (e.g. robins), the specific instances would have been more similar to each other, and, we may imagine, the search would have been somewhat more efficient. We would expect the qualitative pattern of linear search through the visual display and log-linear search through ALTM to remain the same. We turn to the role of object features in the next experiment.

Experiment 3: The interaction of visual guidance and memory search

Figure 1 proposed a simple flowchart model of hybrid search in which linear selection of an object in the visual display cycled with a logarithmic search through memory to determine if the object was in the memory set. Figure 6 presents an elaboration of the model in the light of Experiments 1–2 and in an effort to situate hybrid search amidst other tasks used to study vision, memory, and attention.

Three steps are proposed. First, a visual selection stage in which one (or, perhaps, a small number) of the objects in the visual scene are selected for further processing. Selection is assumed to be serial but “guided” (Wolfe, 1994, 2007). That is, if all targets are known to be red, for example, visual selection will be restricted to red items (Egeth, Virzi, & Garbart, 1984). Second, an identification/categorization stage in which the selected item(s) are compared to the contents of long-term memory and, thus, identified. This stage is assumed to be massively parallel though it does take some time (e.g. Mack, Wong, Gauthier, Tanaka, & Palmeri, 2009). Finally, a stage in which it is determined whether the identified item is a member of the current memory set. The time required for this step is presumed to increase with the log of the memory set size.

Response times for different tasks will depend on different stages in this account. In a standard visual search task, the main driver of RT will be the visual selection stage with the degree of guidance modulating the efficiency of search. In experiments like those on ultra-fast object categorization (Thorpe, Fize, & Marlot, 1996; VanRullen & Thorpe, 2001; Walker, Stafford, & Davis, 2008), or object recognition and categorization more generally (DiCarlo, Zoccolan, & Rust, 2012; Palmeri & Gauthier, 2004), selection is not the issue. These probe the timing of the second stage. In both of these cases, the memory set size can be considered to be one. In classic memory search tasks, visual selection is typically not an issue nor is object identification particularly problematic. Here it is the match to the memory set that is important. Classically, memory search times have been considered to be a linear function of memory set size for memory set sizes within the limits of short term memory (Sternberg, 1966). When the memory set is larger, RT can appear to be logarithmic (Burrows & Okada, 1975) though “memory search” cannot be simply summarized as some simple function of memory set size (Estes, 1988).

Wolfe (2012) found a log relationship between RT and memory set size when the visual set size was one and memory sets could be as large as 100. With a visual set of one, the Hybrid Search task of Wolfe (2012) is, essentially, a memory search task. A central question for the present paper is whether that log function is an idiosyncratic consequence of our massive memory for specific pictures. It would be incorrect to propose that every memory search task will produce logarithmic RT x memory set size functions (Nosofsky, Cox, Cao, & Shiffrin, 2013) but the results of Experiment Two suggest that the logarithmic functions are not limited to the case of specific, photorealistic targets. When the memory set consists of categories of items, RT is still an apparently logarithmic function of the number of such categories, held in memory.

We can describe Experiment 1, within the framework of the Hybrid Search model in Figure 6. Faced with an array of alphanumeric characters, the observer selects one, essentially at random (Box 1: Visual Selection). In Box 2, that item is identified (e.g. as a “7”) and categorized (“a number”). In Experiment 1a, the target category label (e.g. “letter”) constitutes a memory set of 1. Thus, in Box 3, the size of the set of numbers or letters does not matter. In this particular example, “7” is not a letter and the process loops back to select another item. In Experiment 1b, the memory set would be some arbitrary set of items (e.g. “A H 5 7”). In that case, the Box 3 contribution to the RT would be related to the log of the memory set size.

Note that this flow chart does not offer an account of why search for categories is slower than search for objects. Presumably, as noted above, that is related to the relative lack of specificity in the representation of “animal” as compared to “this picture of a rabbit”. However, the details of that difference are beyond the scope of this simple flow chart.

Experiment 3 is designed to further illustrate how these three steps interact to produce hybrid search behavior. Consider the situation illustrated in Figure 7. The observer holds a memory set of 9 animals in mind. The visual array can contain three types of items. An example of each is circled in the figure. Starting from the left of the “visual array”, there are animals. They are candidates for selection (Figure 6: Box 1). They will be identified and categorized as animals (at the superordinate level) (Figure 6: Box 2). Selection of an animal will invoke a memory search to determine if that animal is in the memory set (Figure 6: Box 3). If we vary the number of animals in the visual array, each additional item should incur the relatively high cost of performing all three steps. Moreover, the cost of additional animals in the visual display will be directly related to the number of animals in the memory set because the cost of the memory set step will increase logarithmically as the memory set increases.

In the second circle, an alphanumeric character is an example of a stimulus that should not even be selected. Even though “animal” is a diverse class of visual stimuli, an “F” is never going to be an animal. In the absence of animal-like basic features, guided search will keep attention away from such stimuli. Thus, if we vary the number of letters in this display, there should be essentially no cost of additional letters. In the third circle is an object that might be selected (Figure 6: Box 1). In Box 2 (Figure 6), it would be identified (“boxing gloves”) and categorized (as something other than “animal”). It could then be rejected without a search of the memory set because, not being an animal, it cannot be in an all-animal memory set. This could be seen as “guided” memory search. Additional distractors of this sort in the visual array will produce a cost related to visual selection and/or identification & classification. It will not produce a memory search cost dependent on the size of the memory set.

Experiment 3 tests and supports these predictions. Specifically, in a hybrid search for specific members of a well-defined category, three types of non-targets are important: 1) members of the category, 2) items that share basic visual features with the target set but that, once identified, can be rejected as members of a non-target category, and 3) items that are sufficiently different from targets in their visual features and category that they do not need

to be selected at all in a visual search. As the results of Experiment 3 will show, the RT cost of each additional in-category distractor includes the cost of a memory search through the target set in ALTM. However, the cost of each additional visually similar but categorically incorrect distractor includes only the usual costs of inefficient visual search without a search through ALTM. Finally, there is no cost for each additional dissimilar distractor as these are neither selected in visual search nor do they provoke a search of ALTM.

Methods

In Experiment 3, the stimuli were photographs of real objects. More than 200 images were collected in each of seven categories. The target category was the superordinate category “Animals”. In Experiment 3a, distractors could be drawn from the Animal category or from one of three other non-target, superordinate categories (clothing, musical instruments, and plants). These non-target categories were picked to be categorically different from Animals. However, it was intended that the items in the distractor categories should share low-level features with the target category (e.g. curved contours, irregular shapes, especially when projected into a 2D plane). In 3b, the nontarget categories were flags, money, and picture frames. Here, it was intended that the items in the distractor categories should be clearly distinguished from the target category by low-level features (e.g. straight contours and regular shapes). Items in this set were either rectangular or round in a manner that animals are not. Moreover, within a category, these items were more homogeneous in shape than target items (e.g., almost all coins are round). An animal would “pop-out” in an array of these distractor types and it should be possible to guide attention away from distractors and toward potential animals. Note that the selections of distractor categories that were like animals (3a) and not like animals (3b) were the subjective choices of the experimenters. If we knew more about the set of ‘guiding’ features in visual search, especially form features, the distinction might have been based on a more quantitative measure of those features. As the results will show, even in the absence of such knowledge, the subjective choices of distractor categories produced the predicted results.

In Exp 3a, Os memorized a target set of 2, 4, 8, or 16 animals. Once one of those sets was securely held in ALTM, Os performed a block of 512 trials. On each trial, there was a 50% chance of a target being present. There were 0, 4, or 8 animal distractors and 0, 4, or 8 distractors drawn from the non-target categories of clothing, musical instruments, and plants. Thus, overall visual set size could be 4, 8, 12, or 16. The case of 0 animal and 0 other distractors was not used. There were 64 trials for each of the 8 permitted combinations of numbers of animal and other distractors. Individual distractor items did not repeat within a trial. The four memory set sizes were run in pseudo-random order.

Objects were presented in a jiggled 5 by 5 array. Each object fit in a box of 3.7° by 3.7°. The field as a whole subtended 24.5° by 24.5° at the 57 cm viewing distance.

Twelve Os (12 observers (Mean age=27, Stdev=9.2, 9 females) gave informed consent and were paid \$10/hr to participate in this experiment. All had at least 20/25 vision with correction, all passed the Ishihara Color Test, and all were fluent speakers and readers of English.

In Exp 3b, the same 12 Os memorized a target set of 4 or 16 animals. Again, for each memory set, Os performed a block of 512 visual search trials. On each trial, there was a 50% chance of a target being present. Again, displays consisted of 0, 4, or 8 animal distractors and 0, 4, or 8 non-target category distractors. This time, the non-target categories were flags, money, and picture frames. The design was otherwise the same as in the 3a. Indeed, the experiments were run as a single study with the order of all conditions counterbalanced across observers.

Results

Figure 8 shows the slopes of the functions relating RT to the number of items of a specific distractor type, presented in the visual display. The magnitudes of the slope functions give the cost to the visual search of adding distractors of a specific type. Thus, for example, the circles show the cost of adding a distractor to the display when that distractor comes from the categories: plant, clothing, or instruments. As can be seen, that cost is roughly constant at about 20 msec/item regardless of the number of animals in the memory set. The figures and data analysis show the data averaged from 8 of the 12 observers. We removed four Os with error rates over 12% overall and error rates over 25% in individual cells of the experiment. These four Os averaged 22% miss errors (range 14–37%, with error rates over 50% in some cells). The error rates for the 8 Os shown averaged 8% (range 5–10%). It is not desirable to remove 1/3rd of the observers. However, RT analyses are questionable when error rates are high. The power analysis, presented earlier, suggests that 8 Os are adequate. Most importantly, the pattern of results, shown in Figure 8, remains essentially the same if all 12 Os are included. Supplementary Figure 1 presents the average data for all 12 Os to illustrate this point. For both Exp 3a and 3b, using D' as a measure of accuracy, accuracy decreases with increasing visual set size and memory set size (ANOVA, all $p < 0.005$). There is no interaction between these terms.

Looking first at the effect of animal distractors - distractors from the same superordinate category as the targets, it is clear that there is a cost of each animal distractor and that this cost increases dramatically with the memory set size. We can understand this in the context of the model in Figure 6. To determine if an animal distractor is a target, it must be selected in visual search (Box 1). Its identity and membership in the category “animal” must be determined (Box 2) and then that animal must be checked against the memory set of animals held in ALTM (Box 3). As in Wolfe (2012), this memory set cost increases with log of memory set size; hence the compressive, logarithmic shape of these functions. To assess the statistical significance of this pattern of results, the RTs (not the slopes) were submitted to a 2-way ANOVA with memory set size and target category set size as factors. The main effects of memory set size are significant for present and absent trials for animal distractors in Exp 3a (Present: $F(3,21) = 37.7$ $p < 0.001$, $ges = 0.35$; Exp 3a Absent: $F(3,21) = 29.78$ $p < 0.001$, $ges = 0.22$). The main effects of number of animals in the visual set are also significant (all $p < 0.01$, average $ges = 0.62$).

Importantly, the interaction of the memory set size and target category set size is significant for target present ($F(6,42) = 16.86$, $p < 0.001$, $ges = 0.06$) and absent trials ($F(6,42) = 26.78$, $p < 0.001$, $ges = 0.06$). This reflects the fact that the cost of each additional animal in the

visual display goes up with the number of items in the memory set. Therefore, if an item is an animal, it invokes a search of the animals in ALTM, which takes more time when there are more animals in that set.

In Experiment 3a, when an observer encounters a non-animal distractor, the resulting RT cost differs from the cost of an additional animal distractor. There is a cost of non-animal distractors. The main effect of number of non-target distractors in the visual display is significant ($F(3,21) = 31.05, p < 0.001, \eta^2 = 0.04$). Looking at Figure 8, this reflects the fact that the slope of the RT x non-animal set size function is roughly constant at about 20 msec/item. There is a significant effect of memory set size ($F(3,21) = 39.45, p = 0.0001, \eta^2 = 0.68$). This reflects the fact that the overall RT is slower when the memory set size is larger. However, the interaction of memory set size and number of non-animal visual items is *not* significant ($F(6,42) = 1.18, p = 0.34$). An additional nonanimal item in the visual display does not cause an additional search through the animals in ALTM. This is consistent with the model, sketched in Figure 6, where it is proposed that each non-animal may be subject to visual selection, but, once selected, it can be rejected without further search through the memory set.

In Experiment 3b, the non-animal distractors are coins, flags, and picture frames that are intended to be visually discriminable from animals. Looking first at effect of the animal items in the visual set, the main effects of memory set size and animal items in the visual set are significant for present and absent trials (Present: $F(1,7) = 9.43, p < 0.05, \eta^2 = 0.07$; Absent: $F(1,7) = 12.30, p < 0.05, \eta^2 = 0.03$). The interaction of memory set and the number of animals in the visual set is not significant for target present trials ($F(2,14) = 2.55, p = 0.15$) though it is significant for the absent trials ($F(1,7) = 19, p = 0.003, \eta^2 = 0.02$). As can be seen in Figure 8, the animal items in Experiment 3b produce similar slopes to those in Experiment 3a but the Memory Set Size 4 slope is somewhat elevated, perhaps accounting for the reduced significance of the interaction. The effect of animal distractors is nearly the same in Experiments 3a and 3b.

Of more importance, in this case, are the results for the distractors that are not animals. For target present trials, there is a main effect of the number of non-animal items in the display ($F(2,14) = 7.6, p = 0.02, \eta^2 = 0.03$). Note, however, that this is a small *negative* slope. RTs are actually a bit shorter when the number of non-animal distractors is larger. There is no significant interaction of the number of non-animal items in the display with the memory set size ($F(2,14) = 2.32, p = 0.15, \eta^2 = 0.007$). The absence of the interaction suggests that these items were not selected in visual search and did not provoke a search of ALTM. On the absent trials, the main effect and interaction are significant ($F(1,7) > 6.23, p < 0.05, \eta^2 > 0.01$). Again, the effects are in the “wrong” direction. Slopes of the RT x set size function are less than zero (as sometimes happens in searches where distractors are easily rejected (Bravo & Nakayama, 1992)). Moreover, larger memory set sizes produce lower slopes; again, the “wrong” direction. This may simply reflect a speed-accuracy trade-off, as errors double from 4% to 8% for memory set sizes of 4 and 16 items, respectively ($t(7) = 2.73, p = 0.029$, two-tailed). Thus, these non-animal distractors do not appear to be selected at all and do not require a search of ALTM.

Discussion

The results of Experiment 3 are consistent with the model outlined in Figure 6. An observer holds a memory set in ALTM. Visual selective attention, guided by basic visual features, selects items from the visual display as possible targets. These are identified and categorized. If need be, they are then compared to the memory set. The time required for that comparison is a function of the log of the memory set size, but in Experiment 3 that cost is incurred only for those items that are identified as members of the superordinate category, animals, to which all targets are known to belong in this experiment. Two classes of items in the visual display can be dispensed without the need for a log search through ALTM. If a distractor has clear non-animal visual features, it need never be selected in visual search. It is not a candidate target. Consequently increasing the numbers of these distractors: the money, flags, and picture frames of Experiment 3b, incurs no RT cost. If the basic features of a distractor are not sufficient to reject it as an animal, that distractor may be selected as a possible target. That selection proceeds at an effective rate of about 20 msec/item. However, once selected, the item can be categorized as a non-animal and rejected with no ALTM search. Thus, there is no cost related to the size of the memory set.

In this experiment, the target category of animals is superordinate. Would the results have differed if the target category had been basic (e.g. birds) or subordinate (e.g. hawks)? Categorization at the basic level should be faster (Tanaka, Luu, Weisbrod, & Kiefer, 1999) and memory search would probably be somewhat slower as the members of the target category became more similar to each other. However, the overall pattern, shown in Figure 8, should remain essentially the same. Os would still select items, reject categorically incorrect selections, and perform memory searches only on items identified as members of the target category. Additionally, error rates would probably be higher, and continue to increase as the number of similar basic level category items in memory increases. Memorizing and discriminating a few items within a basic level category from distractors of the same basic level category is do able, even out to multiple categories with 16 exemplars per category (Konkle, Brady, Alvarez, & Oliva, 2010). However, if observers were asked to memorize 100s of items within one basic level category and had to discriminate those items from other distractors of the same category, we know from Konkle, et al. (2010) that performance would decrease.

General Discussion

Many real-world search tasks are hybrids of visual search and memory search. Obviously, we need to remember what it is that we are visually searching for (e.g. our mental grocery list). This point has not received much attention because the memory set size has generally been held at one in most visual search experiments. After systematically varying the memory set size, Wolfe (2012) proposed that an essentially linear visual search process interacted with a logarithmic memory search process. The experiments reported here require an elaboration of that view into something like the three-box account illustrated in Figure 6.

This three-box account should not be taken too literally. In particular, it is probably a mistake to think of the process as three steps, occurring in sequence, with one ending as the next begins. Consider the results of Experiment 1b in which Os are looking for any of

between 1 and 8 alphanumeric targets. The slopes of the RT x set size functions (plotted as outline squares in Figure 3) range from about 25 msec/item at memory set size 1 to about 100 msec/item at memory set size 8. This means that throughput rate (the inverse of the slope) for the entire process ranges from about 10 items/second for memory set size 8 to 40 items/second for memory set size 1. (The rates can change by about 2X depending on your model of search - (Horowitz & Wolfe, 1998)). However, the identification / categorization step takes on the order of 100 msec *at minimum* (Kirchner & Thorpe, 2006). Those numbers do not add up, if we assume a sequential series of steps. It is better to think of the process, cartooned in Figure 6, as a pipeline (Khayat, Spekreijse, & Roelfsema, 2006; Ramamoorthy & Li, 1977) or, even more metaphorically, like a “carwash” (Wolfe, 2003). Items are selected to enter the carwash at some rate. Each item takes some time to be washed. Since that time is longer than the time between selections, it follows that more than one item is being identified and categorized at the same time.

This carwash account would hold that the identification/categorization step is not a rate-limiting step in hybrid search (at least, not with the relatively simple identifications required here). The rate of selection does seem to limit the entire process. The status of the memory search is not clear. Further research is needed to determine if two or more items can be tested for membership in the same memory set at the same time in hybrid search or if guided visual selection can co-occur with memory search.

A central goal of this work was to determine if the logarithmic relationship of RT to memory set size was restricted to memory sets of specific objects in specific poses. The present results show that the log relationship is more general. Experiment 1b shows that it is a better approximation to the data than a linear model for arbitrary sets of alphanumeric characters. Experiment 2 shows that the same is true for arbitrary sets of categories like animals and picture frames. Experiment 3 serves as a replication of Experiment 2 for these purposes. In other work, we have found similar results with lists of words (Boettcher & Wolfe, 2014).

The relationship of this result to the broader memory literature is complicated by differences in methods. All of the hybrid search work uses “consistent mapping” in the language of Schneider and Shiffrin (Schneider & Shiffrin, 1977; Shiffrin & Schneider, 1977). Moreover, the hybrid work largely ignores order effects in memorization and testing though these can have strong impacts on RTs (Monsell, 1978). Priming in visual search is a similar temporal order effect that is not considered here. Observers are faster to respond to a target on the current if it was a target on a recent trial (Lamy, Bar-Anan, Egeth, & Carmel, 2006; Maljkovic & Nakayama, 1994; Wolfe, Horowitz, Kenner, Hyle, & Vasan, 2004). Note that a target will appear more frequently when it is a member of a small memory set than if it is a member of a large memory set; at least, if targets in the visual search are drawn at random from the memory set. This does not diminish the interest in the orderly shape of the RT x memory set functions in hybrid search but it does illustrate that more research will be required to understand the forces that generate those data.

Nosofsky et al. (2013) have started that work with a series of experiments that combines the photorealistic stimuli of Wolfe (2012) with more classic methods of the memory search

literature. They test memory sets from 1 to 16 in consistent mapping, varied mapping, and a condition where all items were new on each trial. They used memorization methods that preserved information about presentation order and they used only a visual set size of 1. Under these conditions, they obtained curvilinear RT x memory set size functions that look similar to those in the present paper and to Burrows and Okada (1975). Their methods produced error rates that would complicate asking if the RT function was logarithmic. However, the combination of errors and RTs is more suitable to modeling and Nosofsky et al. report that the errors and RT distributions are accounted for successfully by a version of an exemplar-familiarity model (Nosofsky & Palmeri, 1997). It remains to be seen how the results of these memory search experiments apply to the hybrid search situation where the visual set size is greater than 1. As discussed in the context of the three-box model of Figure 6, this will depend on where the rate-limiting steps lie in combined visual and memory search. Can more than one visual selection and memory search coexist?

Interference between memory sets is another factor not systematically examined in the hybrid search paradigm. Anecdotally, it is surprisingly easy to perform several hundred searches for one set of N objects and then put those out of mind effectively enough to search for a different set of objects on the next block. Still, it seems clear that there will be some interference effects, as have been found in similar searches for small sets of words, for example (Wickens, Moody, & Dow, 1981; Wickens, Moody, & Vidulich, 1985). One can look for interference effects between and even within blocks in a hybrid search paradigm. For instance, it is possible to hold two sets of items in memory and search for members of one set in one context and the other in another context (akin to looking for vegetables in the produce section and candy in the snack food aisle). In preliminary work (Boettcher, Drew, & Wolfe, 2013), we have found that observers can do such tasks but they are imperfect at restricting search to the relevant subset. They pay a modest cost for distractor “lures” from the currently irrelevant set but they rarely produce false alarm errors to those lures. These interference effects, when explored more fully, will reveal much about the capacity and flexibility of ALTM.

Finally, though the present paper shows that the basic pattern of hybrid search is found with memory sets that consist of multiple categories as well as multiple specific objects, the term “category” is used rather broadly in this work. As with interference effects and temporal order effects, category type is not specifically manipulated in these experiments. Thus, for instance, it may be that some categories have a privileged status in visual search (e.g. animals Drewes, Trommershauser, & Gegenfurtner, 2011). Other distinctions between types of categories may be important for the speed of recognition. Rosch et al. (1976) found that participants were fastest to verify object categories at an intermediate level (e.g., bird vs. other animals) than at more superordinate levels (e.g., animal vs. vehicle) or subordinate levels (e.g., Northern Cardinal vs. other birds). Jolicoeur et al. (1984) referred to the fastest level of categorization as the entry level. It may not be that entry level categories are processed first. Processing at multiple levels may start at the same time with quicker decisions possible for some types of category (Mack & Palmeri, 2011a; Mack et al., 2009).

It would be interesting to determine if the pattern of hybrid search was different for a set of entry-level categories as compared to memory sets of superordinate or subordinate

categories. Again, referring back to the three boxes of Figure 6, this will depend on the degree to which object identification is a rate-limiting step in hybrid search.

To summarize: Many real world search tasks are searches for any member of a list of possible targets. There is variation in the precision with which those targets can be defined (“food”, “meat”, “beef”, “this steak”). While the precise nature of the targets modulates the efficiency of the search, the evidence from this paper suggests that search times will rise linearly as the number of items in the display increases and will rise logarithmically as the number of items in the memory set increases.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank Isabel Gauthier, Rob Nosofsky, Justin Jungé, and Chad Marsolek for the helpful comments and feedback on the manuscript. This work was funded by grants to J.M.W. from Google and NIH (EY017001).

References

- Baddeley A. The episodic buffer: a new component of working memory? *Trends in Cognitive Sciences*. 2000; 4(11):417–423. [PubMed: 11058819]
- Balani AB, Soto D, Humphreys GW. Working memory and target-related distractor effects on visual search. *Memory & Cognition*. 2010; 38(8):1058–1076. [PubMed: 21156870]
- Beck VM, Hollingworth A, Luck SJ. Simultaneous Control of Attention by Multiple Working Memory Representations. *Psychol Sci*. 2011; 23(8):887–898. [PubMed: 22760886]
- Boettcher S, Wolfe JM. Searching for the right word: hybrid visual and memory search for words. 2014 submitted.
- Boettcher SEP, Drew T, Wolfe JM. Hybrid search in context: How to search for vegetables in the produce section and cereal in the cereal aisle. *Visual Cognition*. 2013; 21(6):678–682. [10.1080/13506285.2013.844959](https://doi.org/10.1080/13506285.2013.844959)
- Brady TF, Konkle T, Alvarez GA, Oliva A. Visual long-term memory has a massive storage capacity for object details. *Proc Natl Acad Sci U S A*. 2008; 105(38):14325–14329. [PubMed: 18787113]
- Brainard DH. The Psychophysics Toolbox. *Spatial Vision*. 1997; 10(4):433–436. [10.1163/156856897x00357](https://doi.org/10.1163/156856897x00357) [PubMed: 9176952]
- Bravo M, Nakayama K. The role of attention in different visual search tasks. *Perception and Psychophysics*. 1992; 51:465–472. [PubMed: 1594436]
- Burrows D, Okada R. Memory Retrieval from Long and Short Lists. *Science*. 1975; 188(4192):1031–1033. [PubMed: 17759685]
- Chase, WG.; Simon, HA. The mind’s eye in chess. In: Chase, WG., editor. *Visual information processing*. New York: Academic Press; 1973. p. 215–281.
- Cousineau D. Confidence intervals in within-subject designs: A simpler solution to Loftus and Masson’s method. *Tutorial in Quantitative Methods for Psychology*. 2005; 1(1):42–45.
- Cousineau D, Larochelle S. Visual-Memory search: An integrative perspective. *Psychological Research*. 2004; 69(1):77–105. [PubMed: 14986138]
- Cowan N. Evolving conceptions of memory storage, selective attention, and their mutual constraints within the human information-processing system. *Psychol Bull*. 1988; 104(2):163–191. [PubMed: 3054993]
- Cowan, N. *Attention and Memory: An integrated framework*. New York: Oxford U press; 1995.
- Cowan N. The magical number 4 in short-term memory: a reconsideration of mental storage capacity. *Behav Brain Sci*. 2001; 24(1):87–114. discussion 114–185. [PubMed: 11515286]

- Cunningham CA, Wolfe JM. Lions or tigers or bears: Oh my! Hybrid visual and memory search for categorical targets. *Visual Cognition*. 2012; 20(9):1024–1027.10.1080/13506285.2012.726455
- DiCarlo JJ, Zoccolan D, Rust NC. How Does the Brain Solve Visual Object Recognition? *Neuron*. 2012; 73(3):415–434. doi: <http://dx.doi.org/10.1016/j.neuron.2012.01.010>. [PubMed: 22325196]
- Drewes J, Trommershauser J, Gegenfurtner KR. Parallel visual search and rapid animal detection in natural scenes. *Journal of Vision*. 2011; 11(2)
- Duncan J. Category effects in visual search: A failure to replicate the “ohzero” phenomenon. *Perception and Psychophysics*. 1983; 34(3):221–232. [PubMed: 6646963]
- Egeth H, Jonides J, Wall S. Parallel processing of multielement displays. *Cognitive Psychology*. 1972; 3:674–698.
- Egeth H, Marcus N, Bevan W. Target-set and response-set interaction: Implications for models of human information processing. *Science*. 1972
- Egeth HE, Virzi RA, Garbart H. Searching for conjunctively defined targets. *J Exp Psychol: Human Perception and Performance*. 1984; 10:32–39.
- Ericsson KA, Kintsch W. Long-term working memory. *Psychol Rev*. 1995; 102(2):211–245. [PubMed: 7740089]
- Eriksen BA, Eriksen CW, Hoffman JE. Recognition memory and attentional selection: serial scanning is not enough. *J Exp Psychol Hum Percept Perform*. 1986; 12(4):476–483. [PubMed: 2946804]
- Estes, WD. Human Learning and Memory. In: Atkinson, RC.; Herrnstein, RJ.; Lindzey, G.; Luce, RD., editors. *Stevens' Handbook of Experimental Psychology*. Vol. 2. New York: Wiley-Interscience; 1988. p. 352-415.
- Evans KK, Horowitz TS, Wolfe JM. When Categories Collide. *Psychol Sci*. 2011; 22(6):739–746.10.1177/0956797611407930 [PubMed: 21555522]
- Hamilton JP, Mirkin M, Polk TA. Category-level contributions to the alphanumeric category effect in visual search. *Psychon Bull Rev*. 2006; 13(6):1074–1077. [PubMed: 17484438]
- Harris JR, Shaw ML, Bates M. Visual search in multicharacter arrays with and without gaps. *Perception and Psychophysics*. 1979; 26(1):69–84.
- Hick WE. On the rate of gain of information. *Quarterly Journal of Experimental Psychology*. 1952; 4(1):11–26.10.1080/17470215208416600
- Horowitz TS, Wolfe JM. Visual search has no memory. *Nature*. 1998; 394(Aug 6):575–577. [PubMed: 9707117]
- Jolicoeur P, Gluck MA, Kosslyn SM. Pictures and names: Making the connection. *Cogn Psychol*. 1984; 16(2):243–275. [PubMed: 6734136]
- Khayat PS, Spekreijse H, Roelfsema PR. Attention lights up new object representations before the old ones fade away. *J Neurosci*. 2006; 26(1):138–142. [PubMed: 16399680]
- Kirchner H, Thorpe SJ. Ultra-rapid object detection with saccadic eye movements: visual processing speed revisited. *Vision Res*. 2006; 46(11):1762–1776. [PubMed: 16289663]
- Konkle T, Brady T, Alvarez GA, Oliva A. Scene memory is more detailed than you think: the role of categories in visual long-term memory. *Psychol Sci*. 2010; 21:1551–1556.10.1177/0956797610385359 [PubMed: 20921574]
- Konkle T, Brady TF, Alvarez GA, Oliva A. Conceptual distinctiveness supports detailed visual long-term memory for real-world objects. *Journal of experimental Psychology: general*. 2010; 139(3): 558. [PubMed: 20677899]
- Krueger LE. The category effect in visual search depends on physical rather than conceptual differences. *Perception and Psychophysics*. 1984; 35(6):558–564. [PubMed: 6483559]
- Lamy D, Bar-Anan Y, Egeth HE, Carmel T. Effects of top-down guidance and singleton priming on visual search. *Psych Bulletin & Review*. 2006; 13(2):287–293.
- Luria R, Vogel EK. Visual search demands dictate reliance on working memory storage. *J Neurosci*. 2011; 31(16):6199–6207. 31/16/6199 [pii]. 10.1523/JNEUROSCI.6453-10.2011 [PubMed: 21508243]
- Mack ML, Palmeri TJ. The timing of visual object categorization. *Front Psychol*. 2011a; 2:165.10.3389/fpsyg.2011.00165 [PubMed: 21811480]

- Mack ML, Palmeri TJ. The timing of visual object categorization. *Frontiers in psychology*. 2011b;2. [PubMed: 21738514]
- Mack ML, Wong ACN, Gauthier I, Tanaka JW, Palmeri TJ. Time course of visual object categorization: Fastest does not necessarily mean first. *Vision Research*. 2009; 49(15):1961–1968.10.1016/j.visres.2009.05.005 [PubMed: 19460401]
- Maljkovic V, Nakayama K. Priming of popout: I. Role of features. *Memory & Cognition*. 1994; 22(6): 657–672. [PubMed: 7808275]
- McGaugh JL. Time-dependent processes in memory storage. *Science*. 1966; 153(3742):1351–1358. [PubMed: 5917768]
- Monsell S. Recency, immediate recognition memory, and reaction time. *Cognitive Psychology*. 1978; 10(4):465–501. doi: [http://dx.doi.org/10.1016/0010-0285\(78\)90008-7](http://dx.doi.org/10.1016/0010-0285(78)90008-7).
- Neisser U, Novick R, Lazar R. Searching for ten targets simultaneously. *Percept Mot Skills*. 1963; 17(3):955–961.10.2466/pms.1963.17.3.955 [PubMed: 14085133]
- Nosofsky RM, Cox GE, Cao R, Shiffrin RM. Memory Search with Short and Long Lists Viewed from the Perspective of an Exemplar- Familiarity Model. *J Exp Psychol: Human Learning and Memory*, ms. 2013
- Nosofsky RM, Little DR, Donkin C, Fific M. Short-term memory scanning viewed as exemplar-based categorization. *Psychological review*. 2011; 118(2):280. [PubMed: 21355662]
- Nosofsky RM, Palmeri TJ. An exemplar-based random walk model of speeded classification. *Psychol Rev*. 1997; 104(2):266–300. [PubMed: 9127583]
- Olivers CN, Peters J, Houtkamp R, Roelfsema PR. Different states in visual working memory: when it guides attention and when it does not. *Trends Cogn Sci*. 2011; 15(7):327–334. S1364-6613(11)00085-4 [pii]. 10.1016/j.tics.2011.05.004 [PubMed: 21665518]
- Omohundro J, Homa D. Search for abstracted information. *American Journal of Psychology*. 1981; 94:267–290.10.2307/1422745
- Palmer EM, Fencsik DE, Flusberg SJ, Horowitz TS, Wolfe JM. Signal Detection Evidence for Limited Capacity in Visual Search. *Atten Percept Psychophys*. 2011; 73(8):2413–2424.10.3758/s13414-011-0199-2 [PubMed: 21901574]
- Palmeri TJ, Gauthier I. Visual object understanding. *Nature Reviews Neuroscience*. 2004; 5(4):291–U217.10.1038/nrn1364
- Pelli DG. The VideoToolbox software for visual psychophysics: transforming numbers into movies. *Spatial Vision*. 1997; 10(4):437–442.10.1163/156856897x00366 [PubMed: 9176953]
- Poole BJ, Kane MJ. Working-memory capacity predicts the executive control of visual search among distractors: The influences of sustained and selective attention. *Quarterly Journal of Experimental Psychology*. 2009; 62(7):1430–1454.10.1080/17470210802479329
- Ramamoorthy CV, Li HF. Pipelined Architecture. *Computing Surveys*. 1977; 0(1):61–102.
- Rosch E, Mervis CB, Gray WD, Johnson DM, Boyes-Braem P. Basic objects in natural categories. *Cogn Psychol*. 1976; 8(3):382–439. doi: [http://dx.doi.org/10.1016/0010-0285\(76\)90013-X](http://dx.doi.org/10.1016/0010-0285(76)90013-X).
- Rosenzweig MR, Bennett EL, Colombo PJ, Lee DW, Serrano PA. Short-term, intermediate-term, and long-term memories. *Behavioural Brain Research*. 1993; 57(2):193–198.10.1016/0166-4328(93)90135-d [PubMed: 8117424]
- Schneider DW, Anderson JR. A memory-based model of Hick's law. *Cogn Psychol*. 2011; 62(3):193–222.10.1016/j.cogpsych.2010.11.001 [PubMed: 21293788]
- Schneider W, Shiffrin RM. Controlled and automatic human information processing: I. Detection, search, and attention. *Psychol Rev*. 1977; 84:1–66.
- Shepard RN. Recognition memory for words, sentences, and pictures. *J Verbal Learning and Verbal Behavior*. 1967; 6:156–163.
- Shiffrin MR, Schneider W. Controlled and automatic human information processing: II. Perceptual learning, automatic attending, and a general theory. *Psych Review*. 1977; 84:p127–190.
- Sobel KV, Gerrie M, Poole BJ, Kane MJ. Individual Differences in Working Memory Capacity and Visual Search: The Roles of Top-Down and Bottom-Up Processing. *Psych Bulletin & Review*, ms reviewed 9.10.06. 2006

- Soto D, Heinke D, Humphreys GW, Blanco MJ. Early, involuntary top-down guidance of attention from working memory. *J Exp Psychol Hum Percept Perform.* 2005; 31(2):248–261. [PubMed: 15826228]
- Standing L. Learning 10,000 pictures. *Q J Exp Psychol.* 1973; 25(2):207–222.10.1080/14640747308400340 [PubMed: 4515818]
- Standing L, Conezio J, Haber RN. Perception and memory for pictures: Single trial learning of 2500 visual stimuli. *Psychonomic Science.* 1970; 19:73–74.
- Sternberg S. High-speed scanning in human memory. *Science.* 1966; 153:652–654. [PubMed: 5939936]
- Tanaka J, Luu P, Weisbrod M, Kiefer M. Tracking the time course of object categorization using event-related potentials. *NeuroReport.* 1999; 10(4):829–835. [PubMed: 10208556]
- Tanaka JW, Taylor M. Object categories and expertise: Is the basic level in the eye of the beholder? *Cogn Psychol.* 1991; 23(3):457–482.
- Thorpe S, Fize D, Marlot C. Speed of processing in the human visual system. *Nature.* 1996; 381(6 June):520–552. [PubMed: 8632824]
- Treisman AM, Gelade G. A feature-integration theory of attention. *Cogn Psychol.* 1980; 12(1):97–136.10.1016/0010-0285(80)90005-5 [PubMed: 7351125]
- Tulving, E. Episodic and semantic memory. In: Tulving, E.; Donaldson, editors. *Organization of memory.* New York: Academic Press; 1972. p. 381-403.
- Tulving W, Thompson DM. Encoding specificity and retrieval process in episodic memory. *Psychological Review.* 1973; 80:352–373.
- VanRullen R, Thorpe SJ. Is it a bird? Is it a plane? Ultra-rapid visual categorisation of natural and artificial objects. *Perception.* 2001; 30(6):655–668. [PubMed: 11464555]
- Vickery TJ, King LW, Jiang Y. Setting up the target template in visual search. *J of Vision.* 2005; 5(1): 81–92.
- Walker S, Stafford P, Davis G. Ultra-rapid categorization requires visual attention: Scenes with multiple foreground objects. *J Vis.* 2008; 8(4):21. 21-12. [PubMed: 18484860]
- Wickens DD, Moody MJ, Dow R. The Nature and Timing of the Retrieval Process and of Interference Effects. *J Experimental Psychology: General.* 1981; 110(1):1–20.
- Wickens DD, Moody MJ, Vidulich M. Retrieval Time as a Function of Memory Set Size, Type of Probes, and Interference in Recognition Memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition.* 1985; 11(1):154–164.
- Wolfe J, Horowitz T, Kenner NM, Hyle M, Vasan N. How fast can you change your mind? The speed of top-down guidance in visual search. *Vision Research.* 2004; 44(12):1411–1426. [PubMed: 15066400]
- Wolfe JM. Guided Search 2.0: A revised model of visual search. *Psychonomic Bulletin and Review.* 1994; 1(2):202–238. [PubMed: 24203471]
- Wolfe JM. What do 1,000,000 trials tell us about visual search? *Psychological Science.* 1998; 9(1):33–39.
- Wolfe JM. Moving towards solutions to some enduring controversies in visual search. *Trends Cogn Sci.* 2003; 7(2):70–76. [PubMed: 12584025]
- Wolfe, JM. Guided Search 4.0: Current Progress with a model of visual search. In: Gray, W., editor. *Integrated Models of Cognitive Systems.* New York: Oxford; 2007. p. 99-119.
- Wolfe JM. Saved by a log: how do humans perform hybrid visual and memory search? *Psychol Sci.* 2012; 23(7):698–703.10.1177/0956797612443968 [PubMed: 22623508]
- Wolfe JM, Alvarez GA, Rosenholtz R, Kuzmova YI, Sherman AM. Visual search for arbitrary objects in real scenes. *Atten Percept Psychophys.* 2011; 73(6):1650–1671.10.3758/s13414-011-0153-3 [PubMed: 21671156]
- Wong ACN, Gauthier I. An analysis of letter expertise in a levels-of-categorization framework. *Visual Cognition.* 2007; 15(7):854–879.
- Yang H, Zelinsky GJ. Visual search is guided to categorically-defined targets. *Vision Research.* 2009; 49(16):2095–2103. [PubMed: 19500615]

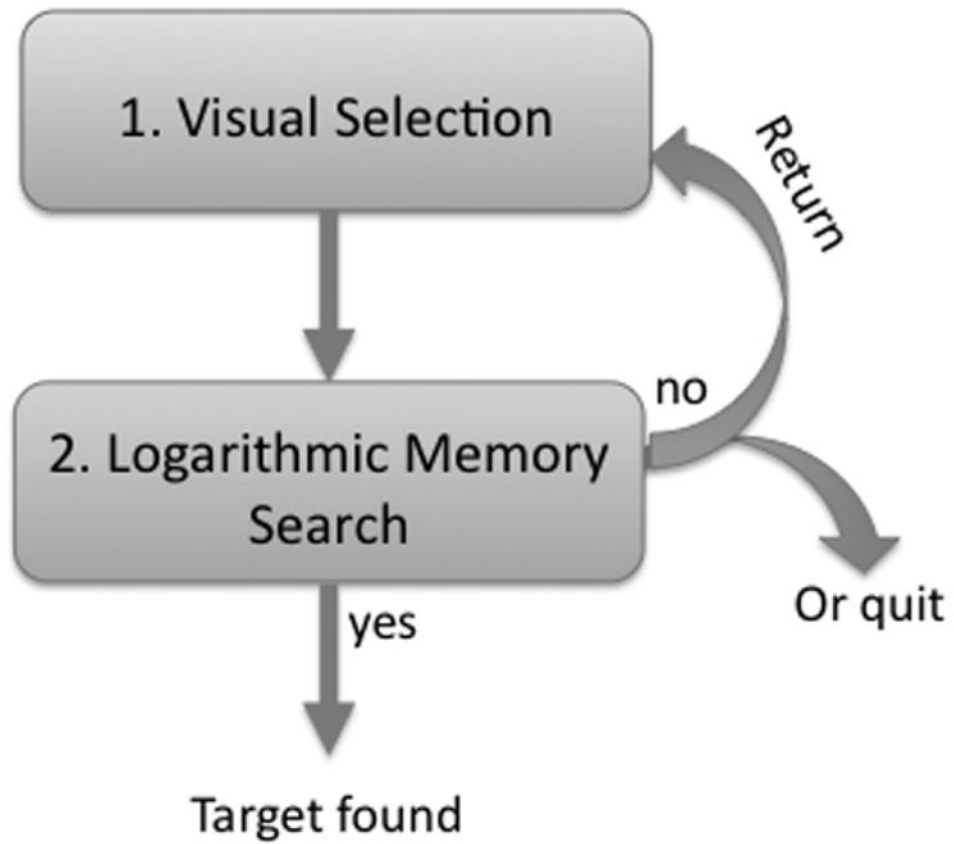


Figure 1.
A simple model of Hybrid Search

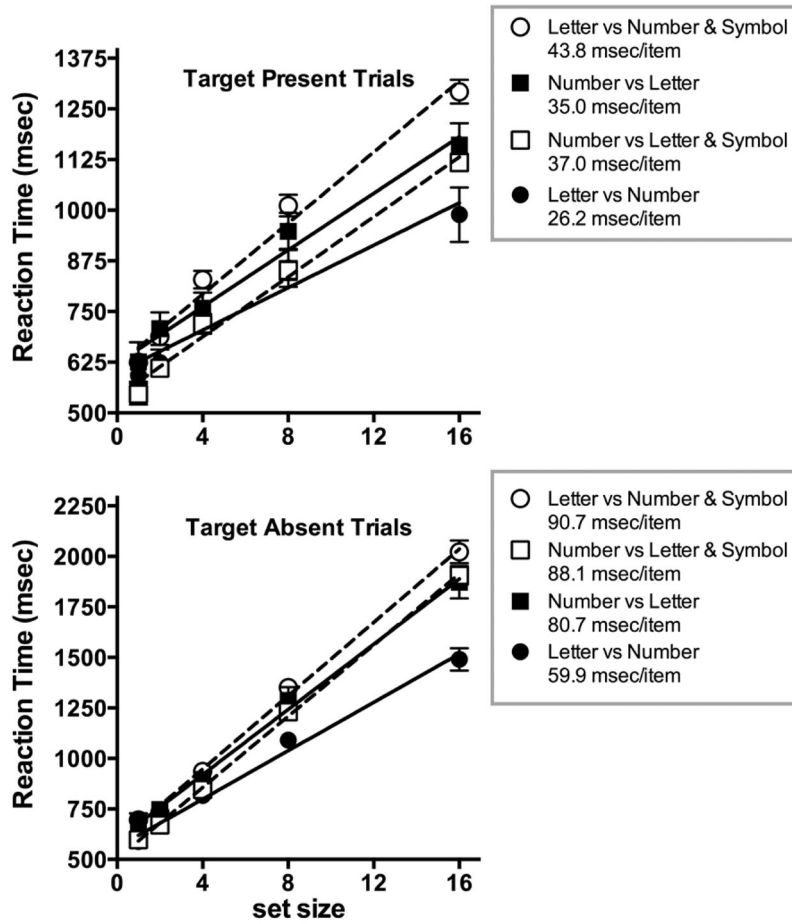


Figure 2. RT x Set Size functions for correct present and absent trials in Experiment 1a. Error bars show ± 1 SEM (within-observer errors, calculated using the method of (Cousineau, 2005)).

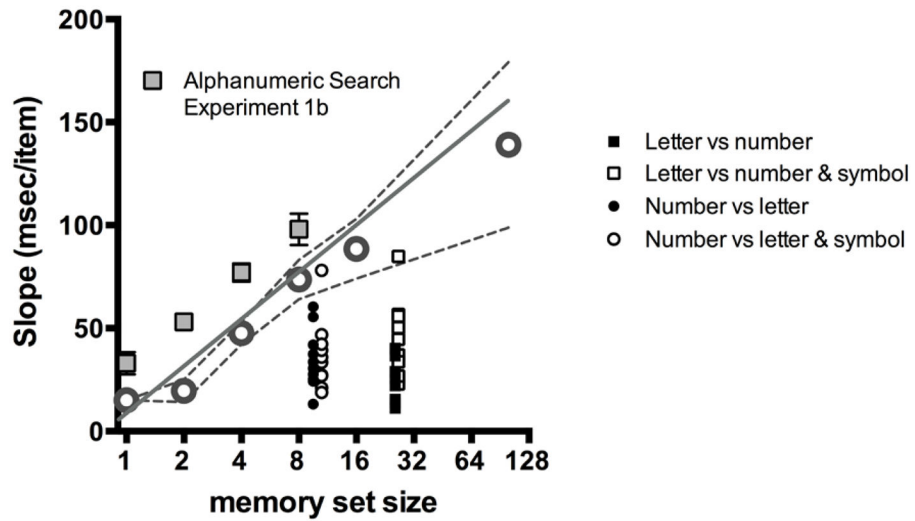


Figure 3.

Results from Experiments 1a & 1b compared to results from (Wolfe, 2012). Data points represent the slope of the RT x Visual Set Size functions as a function of the memory set size on a log scale. Experiment 1a data for individual subjects are shown as black and white small circles and squares. Average Experiment 1b data are shown as outlined squares. Within-subject standard errors fall within the data points. Gray circles are data taken from Wolfe (2012). The line is a best-fit regression. Dashed lines represent 95% confidence intervals around the Wolfe (2012) data.

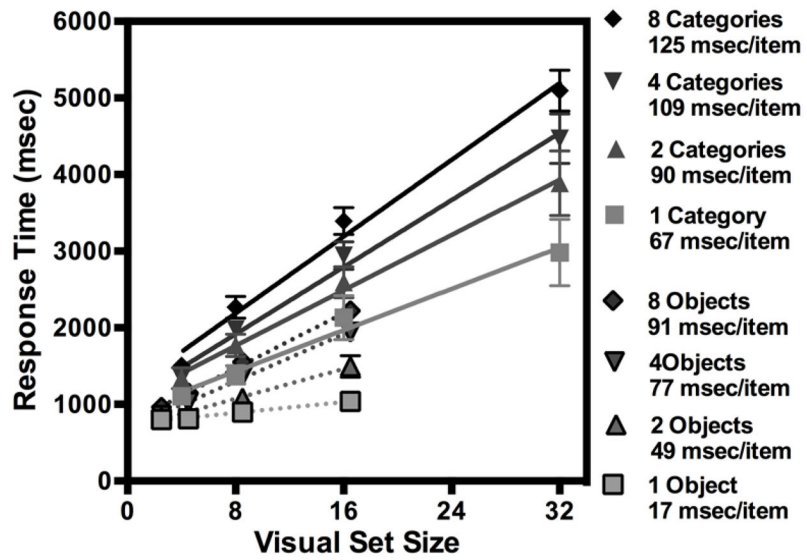


Figure 4. RT as a function of visual set size for Experiment 2 (Solid lines). Data from comparable conditions from (Wolfe, 2012) are plotted with black outline figures and dotted lines (Error bars, where visible, are ± 1 SEM, within observer).

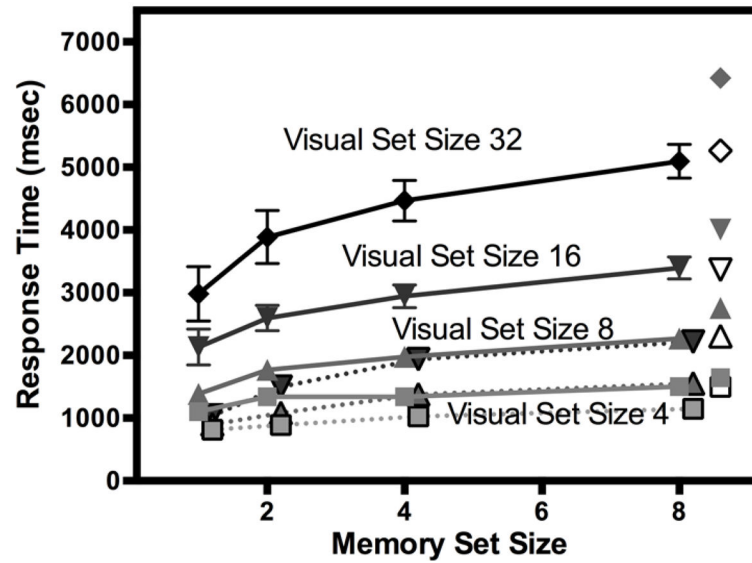


Figure 5.

RT as a function of Memory Set Size. Solid lines show data from Experiment 2. Dotted lines show comparable conditions from (Wolfe, 2012). Symbols to the right show predictions of Memory Set Size 8 data. The shaded figures are based on linear extrapolation from memory set sizes 1, 2, & 4. The open figures are based on logarithmic extrapolation from log₂ (memory set sizes 1, 2, & 4). Error bars, where visible, are ± 1 within observer SEM.

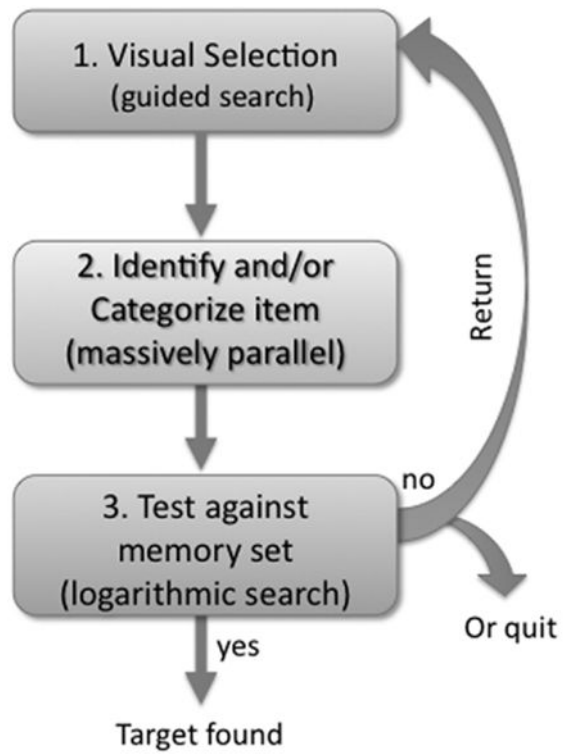


Figure 6.
A more elaborated model of Hybrid Search

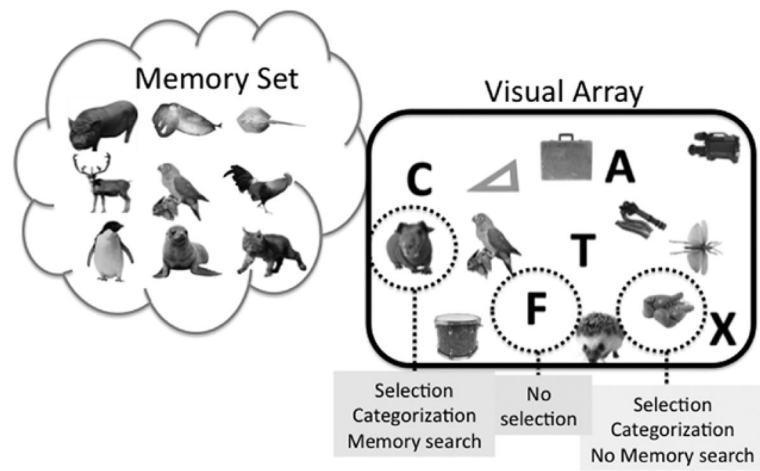


Figure 7.

For a given memory set, there are three types of items in the visual set: Those that could not possibly be targets (e.g. letters, in this case), items that have visual properties like those in the memory set but are categorically incorrect (e.g. boxing gloves are not animals), and those that are both visually and categorically appropriate and, thus, require a search of the memory set to determine their status.

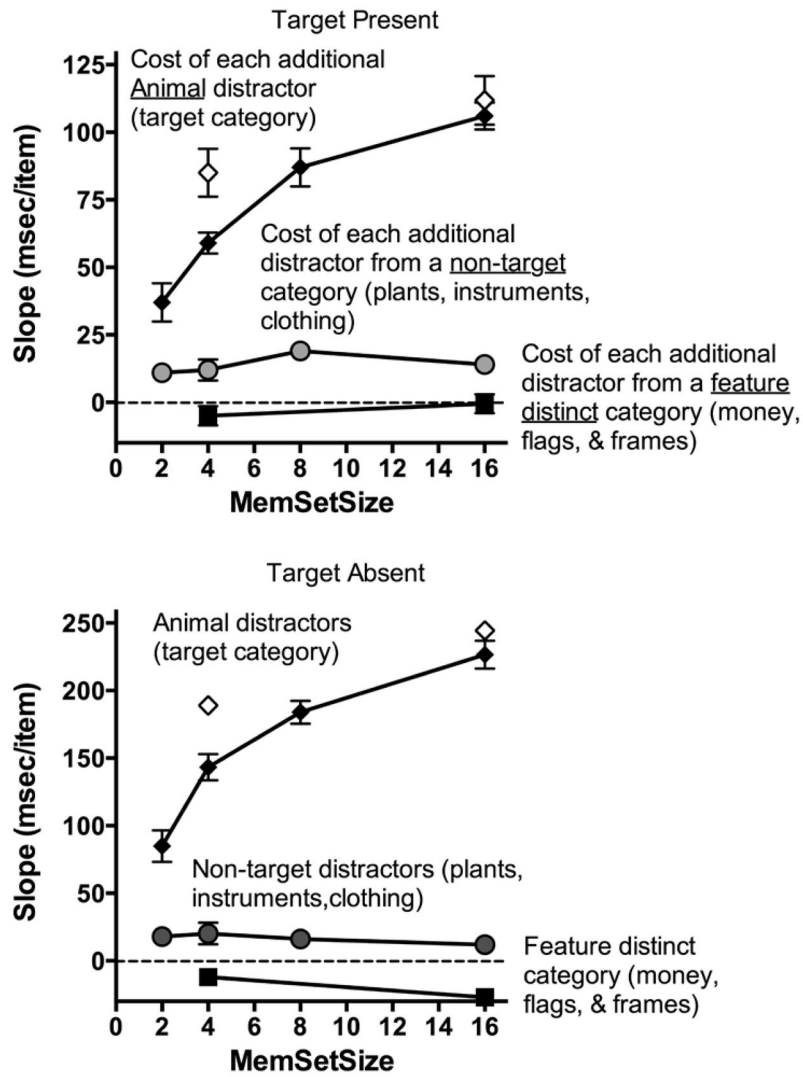


Figure 8.

Slopes of the functions relating RT to number of items in a specific category (target – diamond, closed-Exp 3a, open-3b, non-target – circle, or feature distinct – square) as a function of the memory set size. This is a measure of the cost of each additional distractor of a specific variety. Fig 8a shows average target present for the 8 Os with reasonable error rates. 8b shows target absent trials. Error bars are ± 1 SEM (within-observer errors).