

# Visual search for category sets: Tradeoffs between exploration and memory

**Melissa M. Kibbe**

Department of Psychology and Center for Cognitive  
Science, Rutgers University,  
Piscataway, NJ, USA



**Eileen Kowler**

Department of Psychology and Center for Cognitive  
Science, Rutgers University,  
Piscataway, NJ, USA



Limitations of working memory force a reliance on motor exploration to retrieve forgotten features of the visual array. A category search task was devised to study tradeoffs between exploration and memory in the face of significant cognitive and motor demands. The task required search through arrays of hidden, multi-featured objects to find three belonging to the same category. Location contents were revealed briefly by either a: (1) mouseclick, or (2) saccadic eye movement with or without delays between saccade offset and object appearance. As the complexity of the category rule increased, search favored exploration, with more visits and revisits needed to find the set. As motor costs increased (mouseclick search or oculomotor search with delays) search favored reliance on memory. Application of the model of J. Epelboim and P. Suppes (2001) to the revisits produced an estimate of immediate memory span (M) of about 4–6 objects. Variation in estimates of M across category rules suggested that search was also driven by strategies of transforming the category rule into concrete perceptual hypotheses. The results show that tradeoffs between memory and exploration in a cognitively demanding task are determined by continual and effective monitoring of perceptual load, cognitive demand, decision strategies and motor effort.

Keywords: visual search, active vision, memory, exploration, eye movements, saccades, arm movement, immediate memory, oculomotor, categorization, cognitive load, decision-making

Citation: Kibbe, M. M., & Kowler, E. (2011). Visual search for category sets: Tradeoffs between exploration and memory. *Journal of Vision*, 11(3):14, 1–21, <http://www.journalofvision.org/content/11/3/14>, doi:10.1167/11.3.14.

## Introduction

Active visual tasks, such as searching a room for misplaced keys, or driving a car along an unfamiliar route, can make extraordinary demands on visual, cognitive and motor resources. Information must be gathered from large regions of space and retained for extended periods of time. Visual details are continually being forgotten, and must be retrieved by means of motor actions, such as movements of the eye or head. Decisions about whether to rely on an accumulating (but fragile) memory for the contents of a scene, or to refresh memory by revisiting previously seen locations, may be made at intervals ranging from one to three times each second. These decisions must weigh the risks of relying on a potentially inaccurate memory against the costs in time or effort of generating the motor actions needed to explore the environment. The challenges faced during active tasks increase when the tasks impose significant cognitive requirements involving the generation and evaluation of hypotheses about the contents of the scene. This study investigates the tradeoffs between exploration and memory in a cognitively demanding task that involves both visual search and categorization.

Much recent effort has been devoted to understanding the trade-offs between memory and motor exploration during active tasks. Initial reports emphasized the limited capacity of memory in contrast to the seemingly unlimited ability to generate eye movements (Ballard, Hayhoe, & Pelz, 1995; O'Regan, 1992). This perspective was supported by novel studies of eye movements during “active” visual tasks, showing that people preferred to re-examine previously seen locations, rather than relying on memory, in order to accomplish tasks such as copying arrangements of colored blocks (Ballard et al., 1995) or solving problems in geometry (Epelboim & Suppes, 2001). Subsequent work, however, altered views about the balance between memory and exploration. Studies showed that despite the limits in the capacity of immediate memory for scene details during active tasks, memory can be better than expected, depending on the importance or predictability of the details (Brady, Konkle, Alvarez, & Oliva, 2009; Droll & Hayhoe, 2007; Hollingworth & Henderson, 2002; Pertzov, Avidan, & Zohary, 2009), the location of the details relative to the planned pathway of the saccadic eye movements (Bays & Husain, 2008; Gersch, Kowler, Schnitzer, & Doshier, 2008), or the number of times details were previously viewed (Epelboim et al., 1995; Melcher,

2001; Melcher & Kowler, 2001). In addition, motor exploration proved not to be cost-free. Planning of saccadic eye movements requires time and attention, so that people often avoid making saccades, or decide to alter the saccadic path, if the time needed for planning saccades is too long (Araujo, Kowler, & Pavel, 2001; Coëffé & O'Regan, 1987; Hooge & Erkelens, 1998) or if the distances that must be traveled are large (Ballard et al., 1995; Hardies, Gillner, & Mallot, 2008; Inamdar & Pomplun, 2003). Taken together, these prior findings show that management of resources during active visual tasks is not a matter of favoring either memory or motor planning exclusively, but requires decisions about how to strike the appropriate balance between the two.

The prior work cited above focused mainly on perceptual or perceptual-motor tasks that made significant demands on visual memory and motor planning. Natural tasks, however, often impose significant cognitive demands as well. We investigated the role of both cognitive and motor demands in controlling the tradeoff between memory and exploration by testing performance in a difficult visual search task. The task required searchers to explore arrays of hidden objects to find three multi-featured targets that belonged to the same category. Cognitive demands were controlled by varying the complexity of the rule that defined the category. Motor demands were varied by changing the effector mediating the search (arm or eye) and by imposing different time constraints. The goal was to find out how the cognitive and motor demands of the task affected strategies of relying on memory or exploration.

How might cognitive demands affect strategies for balancing memory and exploration? When thinking and decision making become demanding, the best strategy may be to decrease the reliance on immediate or working memory in favor of a greater reliance on exploration, thereby freeing the limited memory resources for use in thinking or planning, rather than in retaining the contents of the display. Cognitive demands could also have more subtle effects. For example, it is possible that only a subset of the features of fixated objects are encoded during each glance (Alvarez & Cavanagh, 2004; Bays, Catalao, & Husain, 2009; Droll, Hayhoe, Triesch, & Sullivan, 2005; Olson & Jiang, 2002), and the selection of which features to encode may depend on the task demands.

The effect of cognitive task demands on memory has been addressed previously using dual-task methods. In a classic study, Baddeley and Hitch (1974) concluded that cognitive demands of a task do not affect memory based on their finding that words or numbers could be retained in working memory during simultaneous performance of a separate, unrelated task. Other studies have shown that sets of visual objects can be retained in working memory while performing a concurrent visual search task (Woodman, Vogel, & Luck, 2001), although visual search is slowed when the spatial locations of objects have to be retained during search (Oh & Kim, 2004; Woodman & Luck, 2004). Similarly, learning a new, rule-based category was

more difficult when performing a concurrent working memory task, suggesting that working memory resources were required for learning the category (Zeithamova & Maddox, 2006). However, in such dual-task studies, in contrast to most real-world tasks, the to-be-remembered array is unrelated to the primary task. This encourages a compartmentalization of resources in ways that might not be relevant when memory and cognitive demands are integrated into a single task (Sperling & Doshier, 1986), as they are in most natural situations.

### **Present study**

The present study imposed concurrent cognitive, perceptual and motor demands during a visual search task and used the observed search pattern, mediated by either arm or eye movements, to infer how memory resources were managed. The task required subjects to search through an array of hidden, multi-featured objects to find three that belonged to the same category. The task was constructed so that objects could be viewed only one at a time, allowing the searcher to decide when to go back and re-explore a previously visited location. Since only one object was viewed at a time, the experimenters could keep track of these decisions by observing the searcher's motor behavior.

The approach was inspired in part by Epelboim and Suppes's (2001) study of eye movements while solving geometry problems. They used sequences of eye fixations to estimate the span of immediate memory by analyzing the pattern of revisits to previously viewed locations. They estimated the span of immediate memory to be about 4 or 5 regions of a diagram (similar to typical estimates; e.g., Luck & Vogel, 1997), with revisits serving to replenish this limited store as regions were forgotten. We will apply Epelboim and Suppes (2001) model to our search data. (See Zelinsky, Loschky, & Dickinson, 2010, for a similar model of revisits, applied to a memorization task.)

The cognitive demands of Epelboim and Suppes's (2001) geometry task were considerable, and were likely to have played a large role in determining which regions of the diagram were fixated. Subjects decided which regions of a diagram were most relevant as they worked through the problem, and thus were able to make strategic decisions about limited resource allocation "online". This makes for a more natural task in which limited resources must be allocated dynamically. However, because of the complex nature of the geometry problems it is difficult to systematically quantify how task demands affected memory use.

In the present study, cognitive task demands were controlled by varying the complexity of the search rules, and motor demands were controlled by varying the time and effort needed to visit locations. Specifically:

(1) *Variation in the cognitive demands of the task.* The cognitive demands of the search task were manipulated by varying the complexity of the rule that defined the set of target objects. Previous work has shown that the

subjective difficulty of a category rule depends on both the number of features relevant to the rule, as well as on the nature of the decisions about candidate targets (e.g., searching for objects that share one or more features vs. objects that differ on one or more features). Feldman (2000), following on classical work by Shepard, Hovland, and Jenkins (1961), found that the difficulty of learning a new category from examples was proportional to the shortest propositional formula that is logically equivalent to that category; the more logically complex the formula, the more difficult the category was to learn (see also Aitkin & Feldman, 2006; Feldman, 2006; Pothos & Chater, 2002; Pothos & Close, 2008). More recently, Jacob and Hochstein (2008) found that, in a search task in which subjects had to find sets of objects with either the same features or different features, same-feature sets were detected more quickly than different-feature sets. They concluded that detecting similarities results from use of a basic, built-in perceptual process, and is thus less effortful than finding differences.

The present study used five different category rules which were based on either one, two, or three of the objects' four features. Each rule required evaluating either conjunctions of features (each feature value is the same), or exclusive disjunctions of features (each feature value is completely different), or both. Based on prior work (Aitkin & Feldman, 2006; Feldman, 2006; Jacob & Hochstein, 2008), the complexity of the five rules used in the current task was assumed to increase by either adding a feature to the rule or by incorporating a disjunction. If the cognitive demands of the task influence memory use, we expect that as category complexity increases, searchers should visit and revisit more object locations, rather than rely on memory, to make the decision. Two additional aspects of the task should be emphasized. First, the task involves category search, and not category learning, since it required finding a set of objects that satisfied a rule presented before each trial (several possible sets that could satisfy the rule were available in each display). Second, the contents of the displays were chosen so that an "ideal searcher" with perfect memory could find the targets after searching the same number of object locations (between 4 and 5) regardless of the category rule. Thus, an effect of the type of category rule on search would imply that the human searcher (unlike the ideal searcher) was encountering performance limits due to limitations imposed by cognitive processes and strategies and not due to statistical fluctuations within the display.

(2) *Variation in the motor costs of the task.* Motor costs were varied by testing both manual search (Experiment 1), in which subjects searched through objects by clicking locations with a mouse, and the (presumably less demanding) oculomotor search (Experiment 2), in which fixation on a location revealed the object. Two kinds of oculomotor search were tested: *search with delay*, in which a brief pause was imposed between fixating a location and the appearance of the object at that location, and *search with*

*no delay*, in which the object appeared as soon as the fixation was detected. In the delay condition, the duration of the delay was chosen such that the time to carry out the search approximated that of the manual search found in Experiment 1. When the motor demands of a task are high (as in manual search, or in delayed oculomotor search), the best strategy may be to reduce exploration, thus minimizing the amount of time or physical action required for the task. If this were the case, manual search or oculomotor search with delay would be carried out with fewer visits and revisits to object locations than oculomotor search with no delay.

As a preview: Manipulations of the cognitive and motor demands of the task altered the search patterns, and, by implication, the use and reliance on memory. Increasing the cognitive demands of the task, and decreasing the motor demands, each resulted in more visits and more revisits to object locations, that is, a bias to favor exploration over memory. Further analyses done to estimate the span of immediate visual memory from the pattern of revisits using the model of Epelboim and Suppes (2001) yielded estimates similar to those found in prior work with very different task constraints (Epelboim & Suppes, 2001; Jacob & Hochstein, 2009), although the estimates of immediate memory span did vary as a function of both the motor demands and rule complexity. These results show that concurrent monitoring of perceptual states, cognitive load, and motor effort determine the strategies used to control the balance between exploration and memory during active visual tasks.

A portion of these results were presented at meetings of the Vision Sciences Society (Kibbe, 2008; Kibbe, Kowler, & Feldman, 2009).

## Experiment 1

### Methods

#### Subjects

Eight subjects participated. Subjects were either undergraduates recruited from the General Psychology subject pool who earned course credits for participation, paid subjects who earned \$10 for participation, or graduate student volunteers. Subjects all had normal or corrected-to-normal vision. Four subjects completed two sessions of 25 trials, for a total of 50 trials each, and four subjects completed one session of 25 trials. An additional subject was tested but the data were not analyzed due to a visual impairment not disclosed prior to experiment.

#### Stimuli

Stimuli were displayed on a Dell 19" LCD monitor (refresh rate 75 Hz) viewed from a distance of 118 cm.



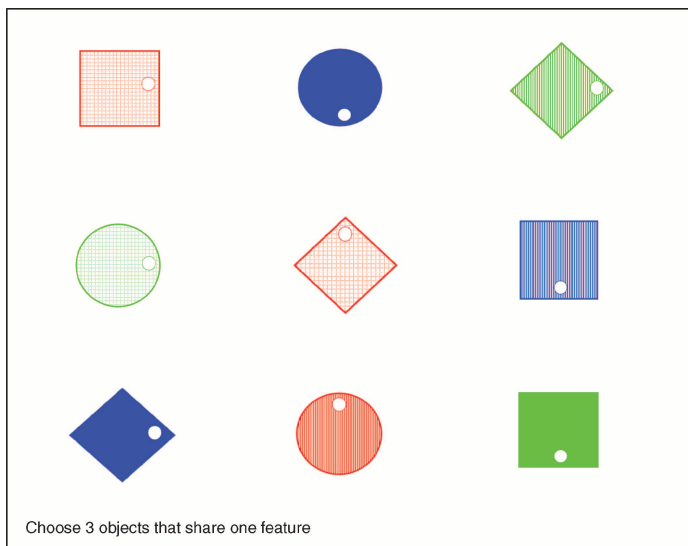


Figure 1. A sample array of objects for a trial. In the actual experiment, these objects were hidden from view and could be revealed one at a time by either a mouse click (Experiment 1) or an eye fixation (Experiment 2). The category rule is displayed at the bottom of the screen at all times. This sample shows Category S (Objects share one feature). There are 10 possible correct sets in the sample array.

Displays consisted of nine “hidden” objects ( $2.9^\circ \times 2.9^\circ$ ) arranged in a  $3 \times 3$  array ( $7.8^\circ$  horizontally by  $7.0^\circ$  vertically). The location of each object was indicated by a black outline ( $3.1^\circ \times 2.8^\circ$ ) and the distance between midpoints of objects was  $4.6^\circ$  horizontally and  $4.4^\circ$  vertically. Objects were revealed by clicking on their location with a mouse. Each object was defined by four trinary features: color (red, green, or blue), shape (oval, rectangle, or diamond), texture (solid, striped, or grid), and orientation (upright, downward, or sideways). Figure 1 shows a sample array.

On each trial subjects searched for a set of three objects belonging to the same category according to one of the five possible category rules. The five category rules were formed by combining conjunction and disjunction rules over the objects’ features (see Figure 2 for examples):

1. *Category S*: Objects share one feature (the other 3 features are irrelevant);
2. *Category SS*: Objects share two features (the other 2 features are irrelevant);
3. *Category SD*: Objects share one feature, differ on one feature, with the remaining 2 features irrelevant;
4. *Category SSD*: Objects share two features, differ on one feature, with the remaining feature irrelevant;
5. *Category SDD*: Objects share one feature, differ on two features, with the remaining feature irrelevant.

Here, “differ” means that each object must have a different value of the feature (e.g. one is red, one is blue, and one is green.) Since the searcher was never told which specific features to search for, the searcher had to decide

which features and objects to explore and of those, which might satisfy the category.

An experimental session consisted of 25 trials organized into five blocks of five trials each. Each category rule (see above) was tested once per block. The order of testing rules within a block was pseudo-randomized with the constraint that the same category rule was never tested twice in succession and no two categories ever appeared in the same order in each block.

The nine objects on each trial were selected such that an ideal searcher (with no memory loss), limited only by statistical fluctuations in the content of the display, would have about the same probability of finding a correct set of three objects for each category rule, regardless of complexity. To create each trial, an algorithm tested every possible 3-object combination of a randomly drawn set of nine objects against a given category rule. The algorithm chose the nine objects such that on each trial there were nine to 12 possible correct sets of three objects for each rule. If the nine objects drawn did not fit the criterion, a new set of nine objects was drawn. To verify the success of this algorithm, an ideal searcher was programmed to perform the search task by choosing locations to visit at random, store the object at each visited location in memory, and then check its memory after every visit to see whether it had found a set of three objects that satisfy the category. The ideal searcher completed 500 trials of each category type (2500 total trials) and performed the about the same

Category rule	Example
Rule S: Objects share one feature	 Each has the same orientation
Rule SS: Objects share two features	 Each has the same color and texture
Rule SD: Objects share one feature, differ on one feature	 Each has the same texture but different shapes
Rule SSD: Objects share two features, differ on one feature	 Each has the same shape and orientation but different colors
Rule SDD: Objects share two features, differ on two features	 Each has the same texture but different colors and orientations

Figure 2. Category rules and examples of each. During training, subjects were presented with three-object sets and were asked to judge whether they were an example of the category rule.

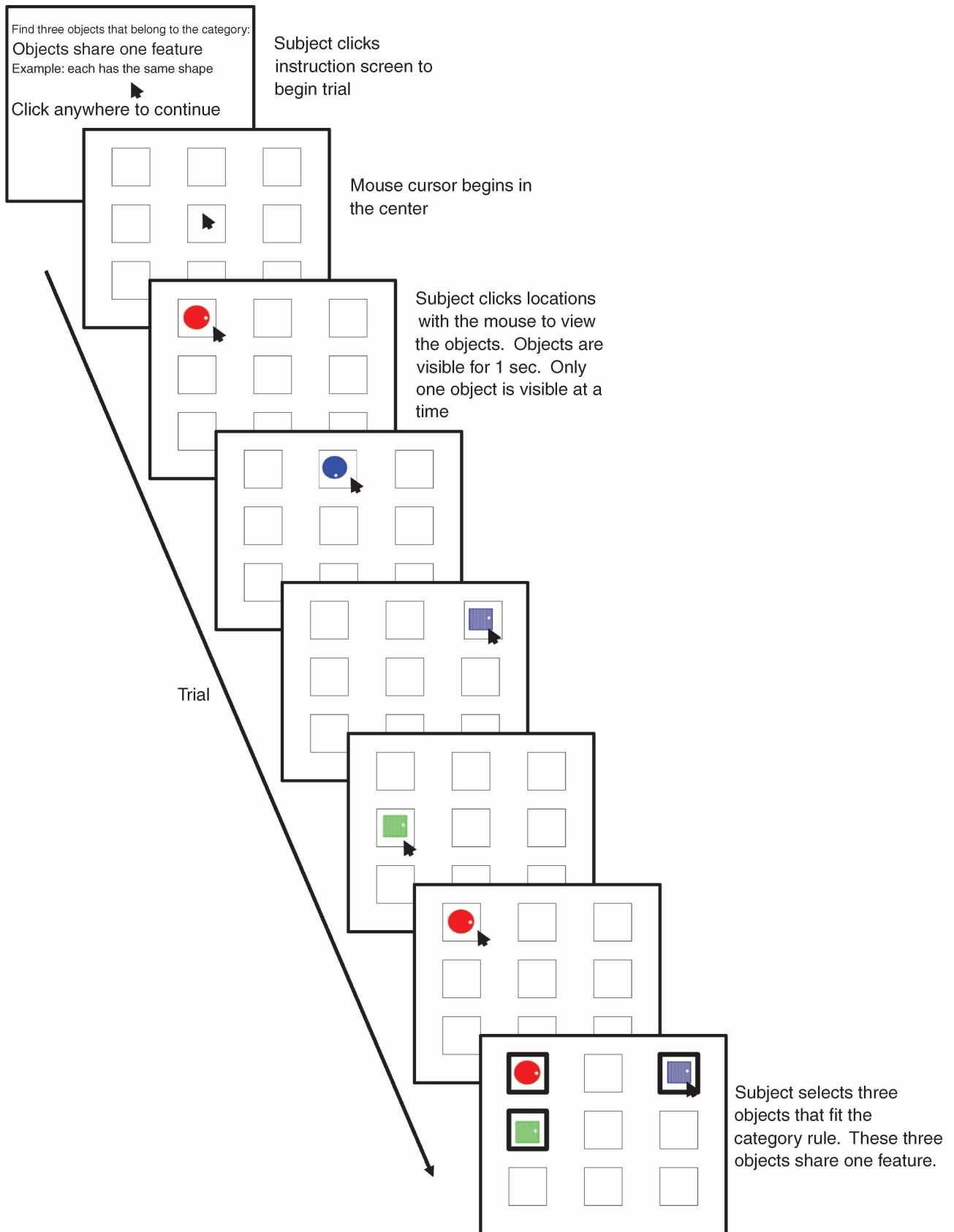


Figure 3. A sample trial for Category S in Experiment 1. Each screen represents the sequence of actions over the course of the trial. Once three objects were selected, the trial ended.

regardless of category rule, requiring an average of 4.7 visits to find a correct set.

### Training

Before beginning the experiment, subjects received 26 training examples to familiarize them with the categories and objects. In each training example, subjects were given a category (e.g. “Objects share one feature”). They were then presented with three objects, and asked to decide whether the three objects belonged to the given category. Subjects were given feedback as to whether they were correct and an explanation as to why or why not. An experimenter observing the subject answered any questions that the subject had about the categories. Subjects were allowed to go through the training examples as many times as they liked until they were confident that they understood the categories. Most subjects felt confident that they had learned the categories after going through all the training examples only once or twice.

### Procedure

A sample sequence of events in a trial are illustrated in Figure 3. Before each trial, an instruction screen appeared that defined the category rule for that trial. When subjects were ready to proceed, they clicked on the instruction screen and nine outline rectangles (one per object) appeared. These nine rectangles indicated the location of each object. The content of a given location was revealed by moving the mouse cursor to the location and clicking on the location. A revealed object remained visible for 1 second. Only one object could be viewed at a time. While one object was visible, clicking on another location had no effect. Inspection of the locations continued until the set of 3 had been found. A right-click of the mouse was used to select a location as belonging to the set. A selected object remained visible and was highlighted with a heavy black border. Once selected, an object could not be unselected. After three objects were selected, the trial ended. It was possible to select one or two objects, and then continue search, however, this strategy was followed only rarely (<1%). (Subjects were advised that selecting only a portion of the set before finding all three objects was likely to lead to error since choices could not be revised.) Subjects were allowed to search for up to two minutes, at which point the trial timed out. No subjects failed to complete the trial in the allotted time. Eye movements were not recorded during in Experiment 1.

### Results

There were three main performance measures: 1) error (selecting a set of objects that did not satisfy the category rule); 2) mean total number of visits to objects per trial;

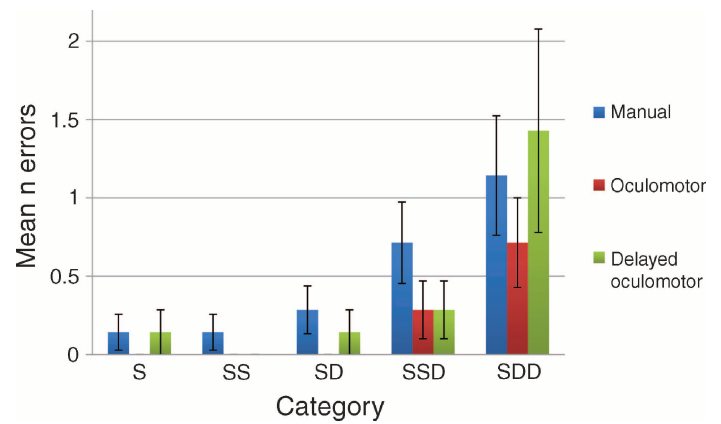


Figure 4. Mean number of errors/subject for each category rule (4–5 trials/category per subject). Error was defined as selecting a set of three objects that did not satisfy the category rule. Error bars represent  $\pm 1$  standard error.

and 3) mean number of revisits to previously viewed objects per trial.

### Error

Error was defined as selecting a set of three objects that did not satisfy the category rule. Figure 4 shows that the mean number of errors within each category rule were low, an average of 1 error or fewer per subject for each category rule, with the highest number of errors for the most complex category. This works out to a total of only 17 trials with errors out of all 275 trials tested (across the 5 rules and 8 subjects). The remaining analyses of visits and revisits are based solely on trials in which an error was not made.

### Visits

The mean number of objects viewed per trial was an indicator of the search strategy. The mean number of visits increased significantly as category complexity increased (Figure 5), from about 9 visits/trial for the easiest category rule, to 17 visits/trial for the most difficult rule ( $F(4) = 7.021, p < 0.001$ ). Post-hoc analysis indicated significant increases in mean visits between adjacent categories, except between Category SD (same on one feature; differing on one feature) and category SSD (same on two features, differing on one feature) (LSD  $p < 0.03$ ).

The large effect of the category rule on the number of visits (Figure 5) was not due to variations in the probability of encountering sets of objects that satisfied the category rule. Displays were constructed so that an ideal searcher, with perfect memory and limited only by the statistical fluctuations in the display, could find a correct set using about the same number of visits across the category rules (see Methods). Results of simulations using the ideal searcher tested in 500 simulations per rule are shown in Figure 5. The ideal searcher found correct sets in

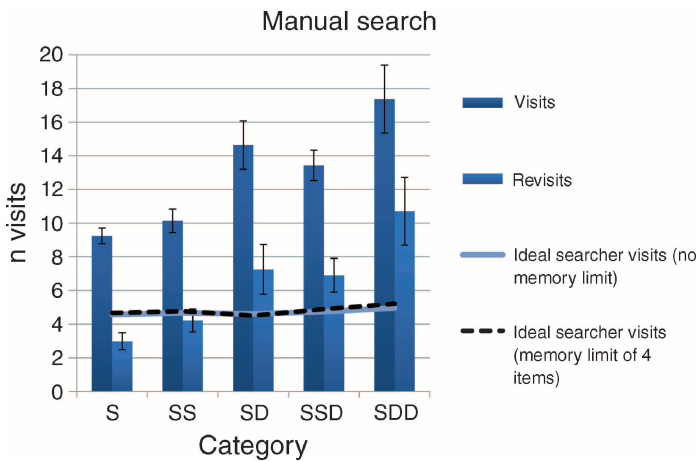


Figure 5. Mean visits and revisits to objects in the display during manual search (Experiment 1). Each bar represents performance averaged over the 8 subjects (subjects were tested in 4 to 5 trials per category rule). Trials on which an error was made were not included. Error bars represent  $\pm 1$  standard error.

an average of 4.7 visits, with small ( $<.5$  object) fluctuations across the rules. Further simulations, in which the ideal searcher's memory was limited to the contents of 4 locations, produced similar results, with 4.8 visits on average required to find the sets, and only small differences in the mean number of objects viewed across the categories. Thus, the effect of rule complexity on the number of visits, shown in Figure 5, represents limitations imposed by cognitive factors or by variation in search strategies across the different rules.

Adding disjunctions to the category rule played a larger role in increasing the number of visits than adding features (see increases between Category SS vs. SD and Category SSD vs. SDD in Figure 5). This suggests that the increases in the number of visits across category rules was not due exclusively to limitations on the number of features that can be held in memory.

### Revisits

Performance was characterized by frequent revisits to previously seen object locations. The mean number of revisits increased as complexity increased ( $F(4) = 9.026$ ,  $p < .001$ , Figure 5), with revisits constituting more than half of the total number of visits for the most complex category. Post-hoc analyses showed that adding a disjunction to the category rule resulted in a significant increase in revisits (Category SS vs. SD: LSD mean difference = 3.27 revisits,  $p < 0.05$ ; Category SSD vs. SDD: LSD mean difference = 5.98 revisits,  $p < 0.001$ .)

### Summary

Search became more difficult as the complexity of the categories increased, requiring more visits to object

locations, more revisits, and resulting in more errors. The effect could not be driven by statistical fluctuations in the stimuli because stimuli were chosen such that an ideal searcher performed nearly identically on each category.

The increase in number of visits across the categories was also not due solely to the number of features defining the category. There were significant increases in the number of visits between categories defined over the same number of features, but differing in categorical structure, as when a category rule contained a disjunction rather than a conjunction. Thus, the effects of category rule on search involved issues of cognitive search strategies, and not exclusively feature memory load.

Experiment 1 required arm movements and mouse clicks to search the object array. Experiment 2 reduced the motor demands by testing search mediated by eye movements. In oculomotor search, the motor demands should be reduced because moving the arm is more effortful and takes longer than moving the eye. In the oculomotor search task, a saccade-contingent method was used so that the contents of each location were not visible until the location was fixated. Given the expected difference between the time needed for oculomotor and manual search, two different types of oculomotor search were tested: 1) a *no-delay* condition in which the contents of a location were revealed as soon as it was fixated, and 2) a *delay* condition in which a brief delay was imposed between the fixation of an object and that object becoming visible.

## Experiment 2

### Methods

#### Subjects

A total of 14 subjects participated, half in the no-delay and half in the delay condition. Subjects had normal or corrected-to-normal vision (soft contact lenses). Two other subjects were excluded prior to the experiment due to failure to acquire a usable signal from the eye tracker because of interference from eyelashes. An additional three subjects were excluded post-experiment due to failure to reach the criterion of choosing a correct set on at least 4 out of 5 trials in each of the 5 categories. Performance below this criterion was regarded as representing failure to understand the task or to make sufficient effort to find a correct set. All subjects completed one block of 25 trials each.

#### Stimuli

Stimuli were identical to those in Experiment 1, except they were displayed on a Viewsonic G90fB 19" CRT



monitor (refresh rate 75 Hz). Movements of the right eye were recorded by an Eyelink 1000 (tower version, with chin and forehead supports). Viewing was monocular.

### Training

Training proceeded exactly the same as in [Experiment 1](#). Eye movements were not measured during training.

### Procedure

The procedure was much the same as [Experiment 1](#). The main difference was that objects were revealed by an eye fixation rather than a mouse click ([Figure 6](#)). The mouse cursor was never visible.

Before beginning the experiment, the standard nine-location calibration incorporated into the Eyelink software was run. This was followed by a test of the gaze-contingent software in which a test stimulus ( $3 \times 3$  array of outline rectangles) was presented. Each location was programmed to be gaze-contingent such that an image appeared in the location only when the location was fixated by the subject. The subject was asked to look around the display and verify that each of the nine test images would appear when fixated. If all appeared when fixated, the experiment proceeded. No subjects failed to meet this criterion. Note that during the experiment fixations that fell between object locations did not result in the appearance of an object.

Each trial consisted of the following events: (1) the instruction screen displaying the category rule for the upcoming trial, which remained on until a button press on the gamepad; (2) the 9-point Eyelink calibration; (3) a repeat of the instruction screen for four seconds; and (4) the appearance of the test stimulus for the trial ( $3 \times 3$  array of nine outline rectangles). In the *No Delay* condition, the content of the locations was revealed by an eye fixation, which was detected online and was visible only while the line of sight remained on the object. When the line of sight fell between objects, which happened occasionally, no object was revealed. Blinks during fixation and fixations between objects had no effect on the visibility of the objects.

In the *Delay* condition a delay was imposed such that an object became visible 750 ms after the fixation was detected rather than immediately upon fixation. The value of 750 ms was chosen based on preliminary testing by independent observers in the lab who judged the duration to be long enough to be noticeable but short enough to not make the task too uncomfortable or unpleasant. During the experiment, no object was revealed if the fixations were too brief, i.e., if the subject looked away from the target within the 750 ms.

Objects were selected as members of the set by fixating the object location and pressing a button on the gamepad. Selected objects remained visible, as in [Experiment 1](#). Usually, subjects selected all three objects at the same

time and did not make further visits. On rare occasions (6% of trials), subjects visited one or two objects after selecting the first two objects in the set but before selecting the third object. If a set was not selected in 2 minutes, the trial automatically timed out.

Subjects were allowed to abort a trial if they experienced difficulty revealing objects via fixations, or if they had mistakenly selected an object, by pressing the trigger on the gamepad. For the no-delay condition, a total of 23 of the 175 trials were aborted (mean per subject: 3.29/25 trials) due to difficulty revealing objects because of loss of signal or drift ( $n = 22$ ) or mistakenly selecting an object the subject did not intend to choose ( $n = 1$ ). For the delay condition, a total of 9 of 175 trials were aborted (mean per subject: 2.8/25 trials) due to difficulty revealing objects ( $n = 5$ ), mistakenly selecting an object ( $n = 2$ ), or accidentally pressing the abort trigger on the gamepad ( $n = 2$ ). An additional 5 trials were aborted automatically because they timed out.

The data reported below were based on the recorded locations of objects that were revealed by a fixation. In some cases subjects fixated a location, looked away at a blank region of the display, and returned to the same location (8% of fixations in the oculomotor (no delay) condition, 4% of fixations in the delayed oculomotor condition). Such sequences were tallied as two consecutive visits in the report of the results below. In addition, the occasional refixation of an already selected object was not counted as a visit in the reported data. Such refixations were rare because in the vast majority of trials (94%) all 3 objects were selected together at the end of the trial.

## Results

### Error

Errors, defined as choosing a set of three objects that did not satisfy the category rule for the given trial, were infrequent (see [Figure 4](#)). Across all subjects, there were a total of 7 trials in which an error was made in 152 completed trials in the no-delay condition, and 14 trials in which an error was made in 161 completed trials in the delay condition. The difference between errors in the delay and no-delay conditions was not significant (Paired  $t = -1.51$ ,  $p = 0.21$ , two-tailed). As in [Experiment 1](#), error increased with category complexity, but never exceeded about 1 trial/condition. The remaining analyses include only correct trials.

### Visits and revisits

Results were similar to [Experiment 1](#) in that the mean number of visits per trial increased across the categories. There was a significant effect of category on the number of visits for both the no-delay ( $F(4) = 6.671$ ,  $p = 0.001$ , [Figure 7](#)) and the delay conditions ( $F(4) = 6.234$ ,  $p = 0.001$ ,



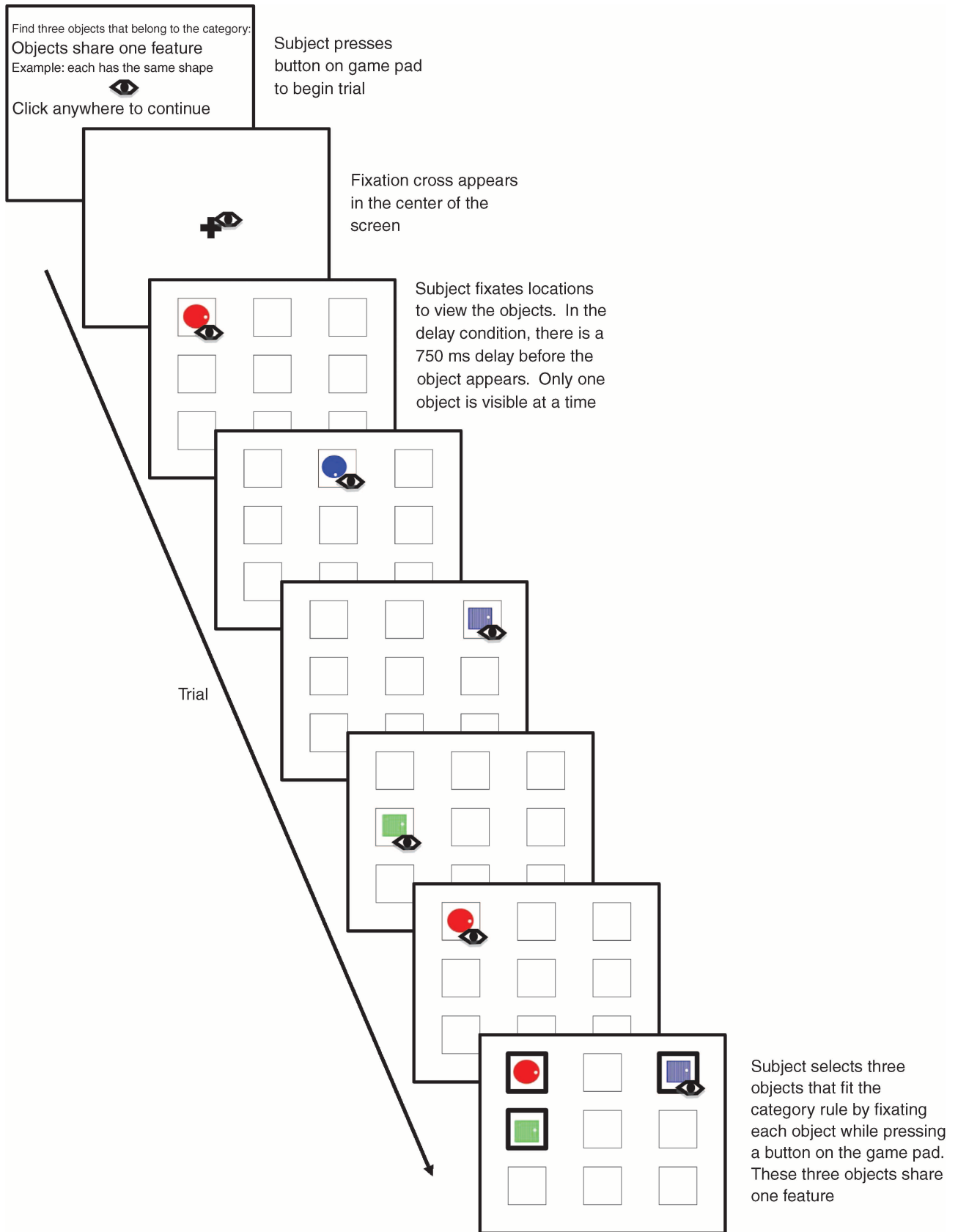


Figure 6. A sample trial for Category S in Experiment 2. Each screen represents the sequence of actions over the course of the trial. Once three objects were selected, the trial ended.

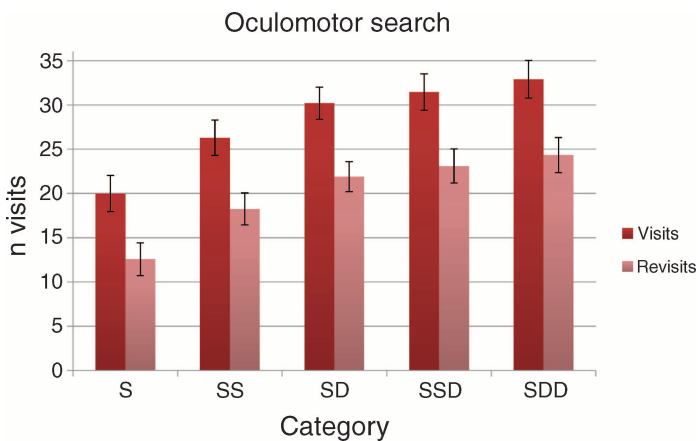


Figure 7. Mean visits and revisits to objects in the display during oculomotor search with no delay (Experiment 2). Each bar represents performance averaged over the 7 subjects (subjects were tested in 4 to 5 trials per category rule). Trials on which an error was made were not included. Error bars represent  $\pm 1$  standard error.

Figure 8). The mean number of times a viewed object was revisited also increased as category complexity increased for both the delay and the no-delay conditions (no-delay condition:  $F(4) = 6.491$ ,  $p = 0.001$ , Figure 7; delay condition:  $F(4) = 7.037$ ,  $p < 0.001$ , Figure 7). Post-hoc analysis of pairwise comparisons between adjacent categories indicated a significant increase in visits between Categories SS and SD for the no-delay condition (LSD  $p = 0.038$ ), and between Categories S and SS for the delay conditions (LSD  $p = 0.034$ ). For the revisits, there was a significant increase Categories S and SS for both delay and no-delay conditions (LSD  $p < 0.05$ ). In the case of oculomotor search, adding both features and disjunctions (and not just adding disjunctions, as was found for manual search) led to more visits and revisits.

Imposing a delay between the landing of the saccade and the appearance of the contents of the location had large effects on performance. The delay increased the total time per trial spent searching (mean = 42.6 s/trial for delay; 28.2 s/trial for no-delay). Nevertheless, despite longer search times in the delay condition, there were about half as many total visits/trial in the delay condition (Figure 8) than in the no-delay condition (Figure 7). The delay also resulted in an average of 65% fewer revisits than in the no-delay condition (Figures 7 and 8). Thus, the delay encouraged a strategy of responding on the basis of fewer total views, a strategy that did not lead to more errors (Figure 4).

Imposing the delay also affected the search rate. Even after accounting for the time consumed by the delay itself, subjects searched about twice as slowly in the delay condition (mean = 2.2 s/object) as in the no-delay condition (mean = 1.2 s/object). These results show that when cost (in time) of search was low, as in the no-delay condition, subjects preferred to visit more objects, and take less time

viewing each one. When cost of search was increased by adding a delay, subjects changed strategy, spending more time viewing each object and visiting fewer objects.

The pattern of performance found for the delayed oculomotor search was similar to that found for manual search in Experiment 1. Specifically, both manual and delayed oculomotor search produced significantly fewer visits and revisits than no-delay oculomotor search (visits:  $F(2) = 175.57$ ,  $p < 0.001$ ; revisits:  $F(2) = 183.26$ ,  $p < 0.001$ ), and there was no difference in the number of visits or revisits between manual and delayed oculomotor search (LSD  $p = \text{n.s.}$ ). There was no significant interaction between search method and category ( $F(8) = 1.619$ ,  $p = \text{n.s.}$ ).

### Summary

The number of visits and revisits to objects increased across categories, regardless of whether search was mediated by arm movements or by saccadic eye movements. Thus, increasing the cognitive demands of the search by adding either features or disjunctions to the category rule encouraged a strategy of increased exploration. In addition, searches with greater motor demands (manual search in Experiment 1, or delayed oculomotor search in Experiment 2) resulted in fewer visits and revisits, that is, less reliance on exploration, than oculomotor search with no delay. The observed number of visits and revisits during delayed oculomotor search was more similar to manual search than to oculomotor search with no delays (see above for statistical support). This suggests that time, rather than motor effort, was the critical factor influencing strategy: faster searches (oculomotor search with no delay) encouraged more exploration.

The effects of both the category rule and delay show that the decision of how to trade off memory and exploration is

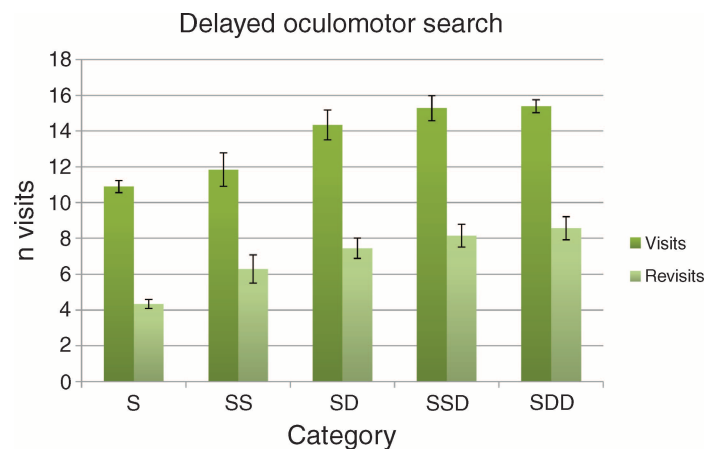


Figure 8. Mean visits and revisits to objects in the display during delayed oculomotor search (Experiment 2). Each bar represents performance averaged over the 7 subjects (subjects were tested in 4 to 5 trials per category rule). Trials on which an error was made were not included. Error bars represent  $\pm 1$  standard error.

affected by both cognitive demands and motor costs. The following analysis uses the pattern of exploration, in particular, revisits to the same object location, to infer how memory was used during search.

### Analysis of the patterns of revisits

The analyses thus far have been focused on the average number of visits and revisits across categories. Another informative aspect of performance is the pattern of revisits to previously viewed object locations (Ballard et al., 1995; Droll & Hayhoe, 2007). Patterns of revisits have been used to infer the amount of information held in immediate, or working, memory during active tasks (Epelboim & Suppes, 2001; Jacob & Hochstein, 2009), and can thus provide insights about the decisions to tradeoff memory and motor activity.

For the current task, the model of Epelboim and Suppes (2001) was fit to the data in an attempt to estimate the span of immediate memory (as they did) and, more generally, characterize the role of memory in the task. Epelboim and Suppes analyzed patterns of eye movements while solving geometry problems to estimate the number of regions of a geometric diagram that are held in immediate memory during the task. In their Oculomotor Geometric Reasoning Engine (OGRE) model, visual working memory is an unordered store of images, the size of which is constant for a given subject and problem. A brief description of the model follows (see Epelboim & Suppes, 2001, for details).

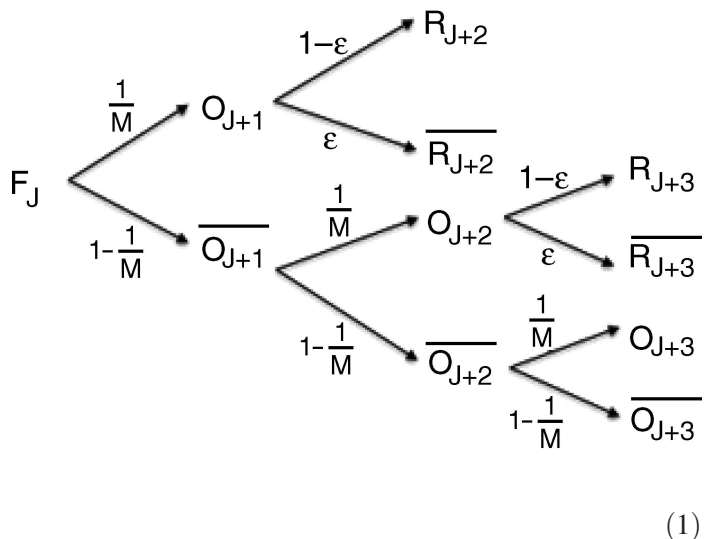
According to the OGRE model, fixation of a region results in the image of that region, denoted  $I(g)$ , to be added to a limited-capacity immediate memory. Regions are added to memory until it reaches its capacity (denoted  $M$ ). At that point, scanning of each new region results in the image of an already-stored region being overwritten. All stored images have an equal probability of being overwritten.

Epelboim and Suppes (2001) stressed that their conception of visual memory is analogous to a mental workspace (e.g., Baddeley & Hitch, 1974), where all the contents are in active use for solving the problem. Therefore, should an image in memory be overwritten, there is a high likelihood that the region will have to be rescanned and added back to visual memory on the next fixation. Specifically, if image  $I(g)$  is overwritten on fixation  $F_{j+1}$ , then the probability that it will be viewed again on the next fixation,  $F_{j+2}$ , is  $1 - \varepsilon$ . This means that, with probability  $\varepsilon$ , an overwritten item will not be rescanned, probably because it is no longer needed for the current cognitive computations. It may be rescanned at a later time if needed again.

Epelboim and Suppes showed that, with probability  $1 - \varepsilon$ , fixation  $F_j$  depends only on the immediately prior fixation,  $F_{j-1}$ , and the state of visual memory immediately prior to the scan,  $V_{j-1}$ . Thus, the path of fixations constitutes a

first-order Markov process, with only rare events (with probability  $\varepsilon$ ) dependent on the distant past.

The procedure used to estimate the size of immediate memory ( $M$ ) from the observed pattern of fixations was derived from the tree below (Epelboim & Suppes, eq. 3), where  $O_j$  denotes overwriting of a region in memory (with probability  $1/M$ ), and  $R_j$  denotes a refixation of that same region on the next fixation, with probability  $1 - \varepsilon$  (bars indicate negation).



The theoretical distribution of re-fixation times,  $k$  (where  $k$  is the number of fixations between the original fixation on a region and a re-fixation) is given by:

$$P(R_{k+J} | O_{k-1+J}, \overline{R}_{k-1+J}, \overline{O}_{k-2+J}, \dots, \overline{R}_{J+1}, \overline{R}_J) = \left(1 - \frac{1}{M}\right)^{k-2} \left(\frac{1}{M}\right) (1 - \varepsilon), k \geq 2. \tag{2}$$

The value of  $M$  that produces the best fit of Equation (2) to the observed distribution of refixation times is the estimated size of visual memory for the given task.

### Applying OGRE to the category search task

Epelboim and Suppes suggested that the OGRE model could be applied to different cognitive tasks requiring the use of visual information. The present search task seemed a promising candidate task for at least two reasons. First, the model assumes that the contents of visual memory are being used for the momentary mental operations required for solving the problem. This means that if an item in memory that is currently in use were to be overwritten, it would be revisited immediately. Since the current task used hidden visual objects, maintaining memory representations of previously viewed visual elements is essential to successful completion of the task. Second, the OGRE

Category	1 <sup>st</sup> order		2 <sup>nd</sup> order	
	Chi <sup>2</sup>	p<	Chi <sup>2</sup>	p<
	Manual search			
S	5.99	.0001	0.07	ns
SS	7.62	.0001	0.22	.025
SD	11.09	.0001	0.15	.001
SSD	6.61	.0001	4.75	.01
SDD	7.60	.0001	0.33	.0001
	Oculomotor Search (no delay)			
S	7.29	.0001	1.14	.0001
SS	11.71	.0001	0.29	.001
SD	15.71	.0001	0.86	.0001
SSD	15.95	.0001	0.48	.0001
SDD	11.53	.0001	0.45	.0001
	Delayed Oculomotor Search			
S	5.17	.0001	0.25	.01
SS	8.65	.0001	0.37	ns
SD	7.97	.0001	0.09	.01
SSD	6.46	.0001	1.52	ns
SDD	7.69	.0001	0.29	.05

Table 1. Results of the Chi Square Test for Independence which tested whether the data from both Experiment 1 (Manual Search) and Experiment 2 (Oculomotor Search and Delayed Oculomotor Search) met the criteria for independence of path outlined by the OGRE model (Epelboim & Suppes, 2001), i.e., the sequence of visits was a first-order Markov process, such that the location searched on visit  $n$  was statistically dependent on the location searched on visit  $n - 1$  and not on the location searched in the earlier visit,  $n - 2$ .

model assumes that the effective size of visual memory can vary from problem to problem. In the current task, the complexity of the categorization rule, as well as the motor demands, were both varied. The OGRE model is thus a promising means of determining whether these variations changed the effective size of working memory ( $M$ ). It is important to stress, however, that  $M$  is an estimate of the number of items held in working memory during the task, not an estimate of memory capacity.

There are also differences between the current search task and Epelboim and Suppes’s geometry problems for which OGRE was proposed. For example: (1) in the category search task, each object contains multiple features, all or some of which may be selectively remembered during a given fixation, while  $M$  is defined in the model in terms of objects (or locations), but not component features, and (2) the search task provides the option of abandoning old object locations and choosing to visit new locations in an attempt to find three objects that fit the category rule, while in geometry the relevance of various portions of the diagram remained fixed as they were constrained by the problem. The implications of these aspects of the category search task will be discussed at the end of this section.

Epelboim and Suppes noted that  $M$  in their study could vary with both subject or with the type of problem. We examined the effect of category complexity and motor costs on  $M$ , with data pooled across subjects.

Before attempting to fit the model, we examined whether the data from both Experiments 1 and 2 met the criteria for independence of path, i.e., the sequence of visits was a first-order Markov process, such that the location searched on visit  $n$  was statistically dependent on the location searched on visit  $n - 1$  and not on the location searched in the earlier visit,  $n - 2$ . Independence of path was tested for each category rule in Experiment 1 and in both conditions of Experiment 2 (delay and no-delay), using a Chi<sup>2</sup> Test for Independence. Results of the tests are shown in Table 1. A strong relationship was found between the locations searched on visit  $n$  and visit  $n - 1$  ( $p < 0.0001$  for all categories and conditions, Table 1). We also found small but significant dependencies between the locations searched on visit  $n$  and visit  $n - 2$  for many of the rules. Although this indicates a higher-order Markov process could be appropriate, we decided that given the strong dependencies between  $n$  and  $n - 1$  relative to the weak dependencies between  $n$  and  $n - 2$  it was reasonable to begin by testing the first-order model.

The span of visual memory,  $M$ , was estimated using the same procedure as Epelboim and Suppes, namely, fitting Equation (2) to the distributions of  $k$ , the number of visits

Category	$M$	Chi <sup>2</sup>
	Manual search	
S	4.58	0.13
SS	4.49	0.03
SD	4.89	0.06
SSD	3.95	0.17
SDD	4.67	0.14
	Oculomotor Search (no delay)	
S	5.28	0.09
SS	5.19	0.16
SD	5.39	0.07
SSD	5.39	0.19
SDD	6.44	0.29
	Delayed Oculomotor Search	
S	5.47	0.24
SS	4.17	0.12
SD	4.94	0.11
SSD	4.34	0.05
SDD	5.09	0.16

Table 2. Estimates of  $M$ , obtained by fitting the OGRE model (Equation 2, see text) to the distribution of  $k$ , the number of locations visited between visiting and revisiting the same location.  $M$  is taken to be an estimate of immediate memory span. Chi<sup>2</sup> values indicate the goodness of the fit of the model to the data, where a value closer to 0 indicates a better fit (all  $p = n.s.$ ).



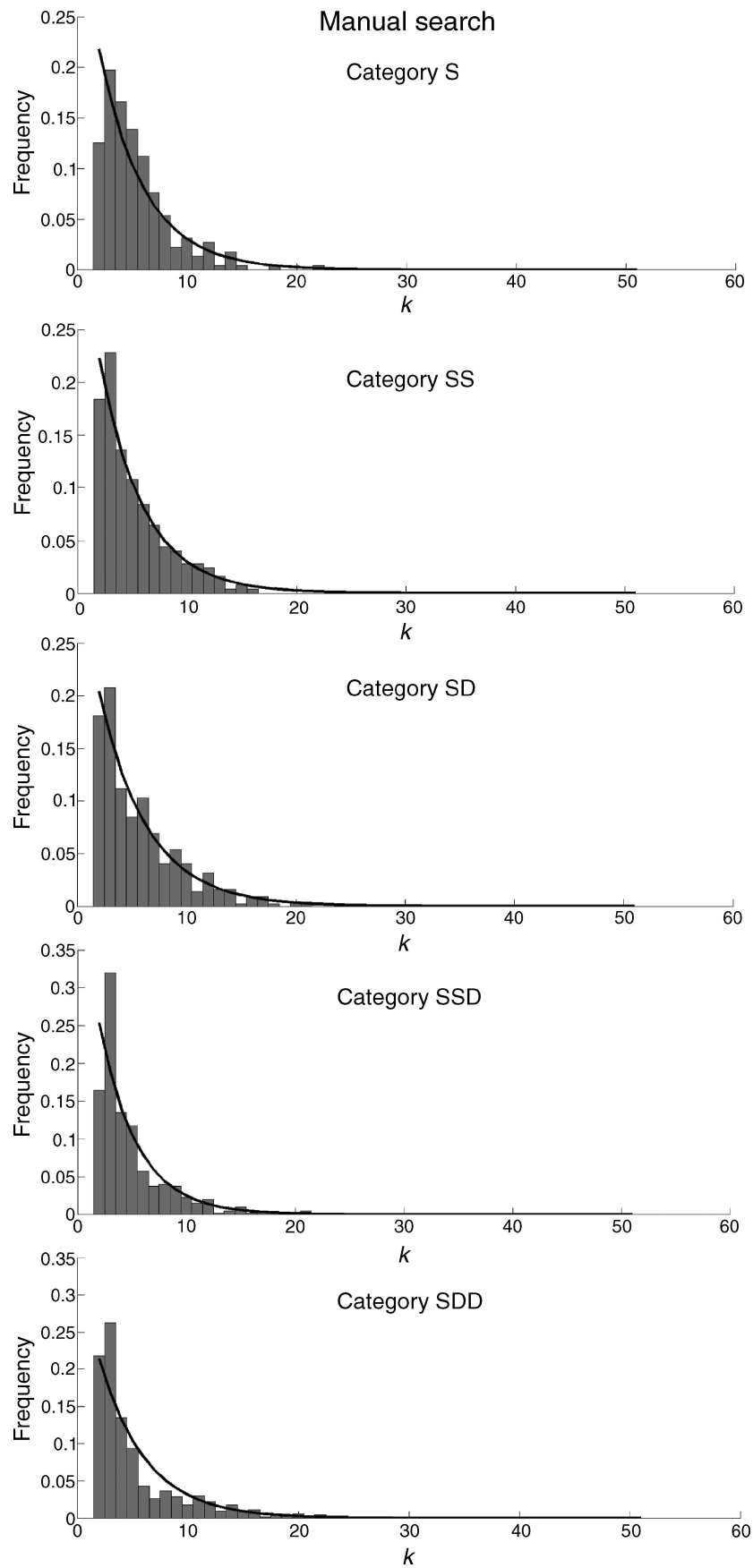


Figure 9. Histograms showing the distributions of  $k$ , the number of locations visited between visiting and revisiting a location, under manual search, [Experiment 1](#). [Equation 2](#) is plotted with the best-fit parameters of  $M$  and  $\varepsilon$  for each category (see text for details).

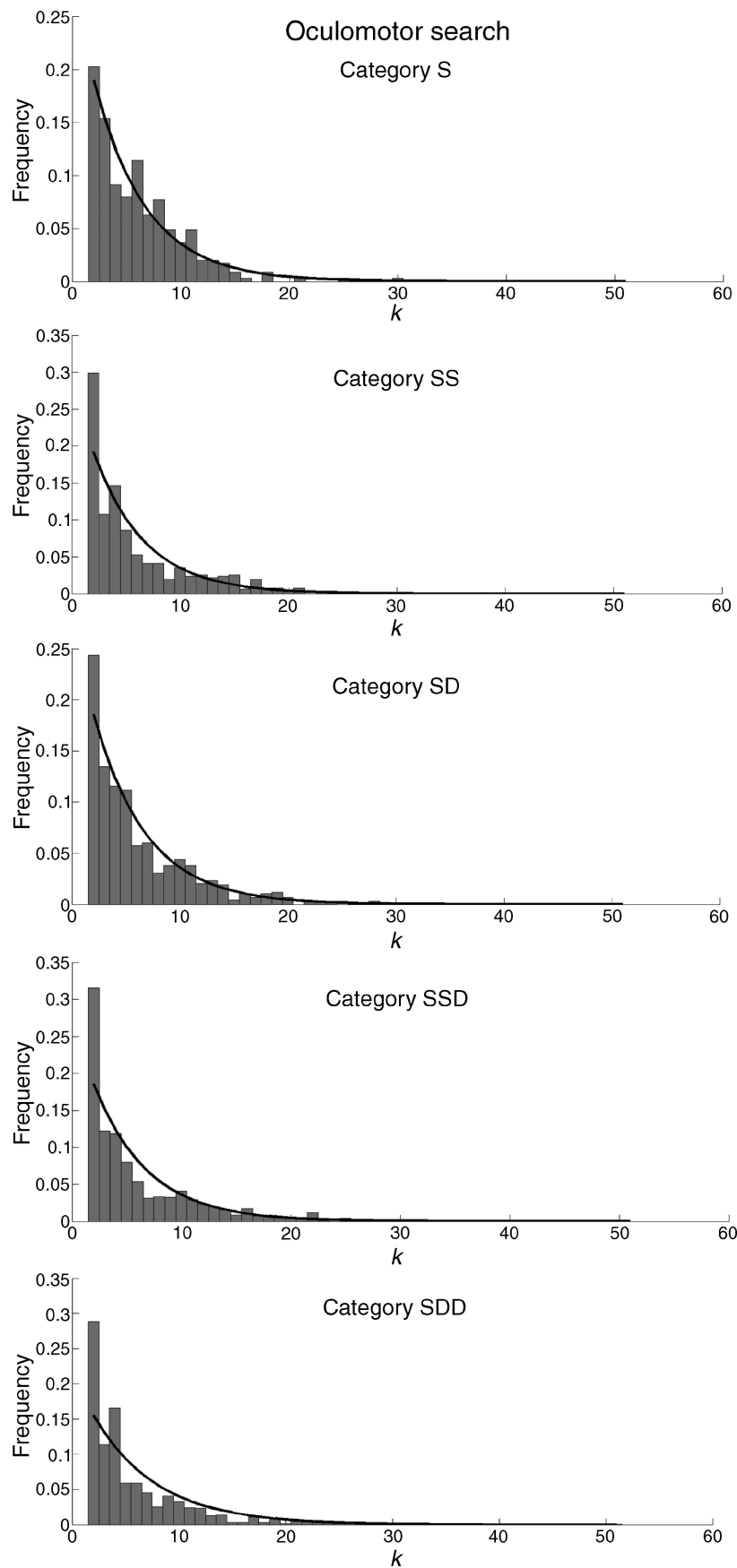


Figure 10. Histograms showing the distributions of  $k$ , the number of locations visited between visiting and revisiting a location, under oculomotor search with no delay, Experiment 2. Equation 2 is plotted with the best-fit parameters of  $M$  and  $\epsilon$  for each category (see text for details).

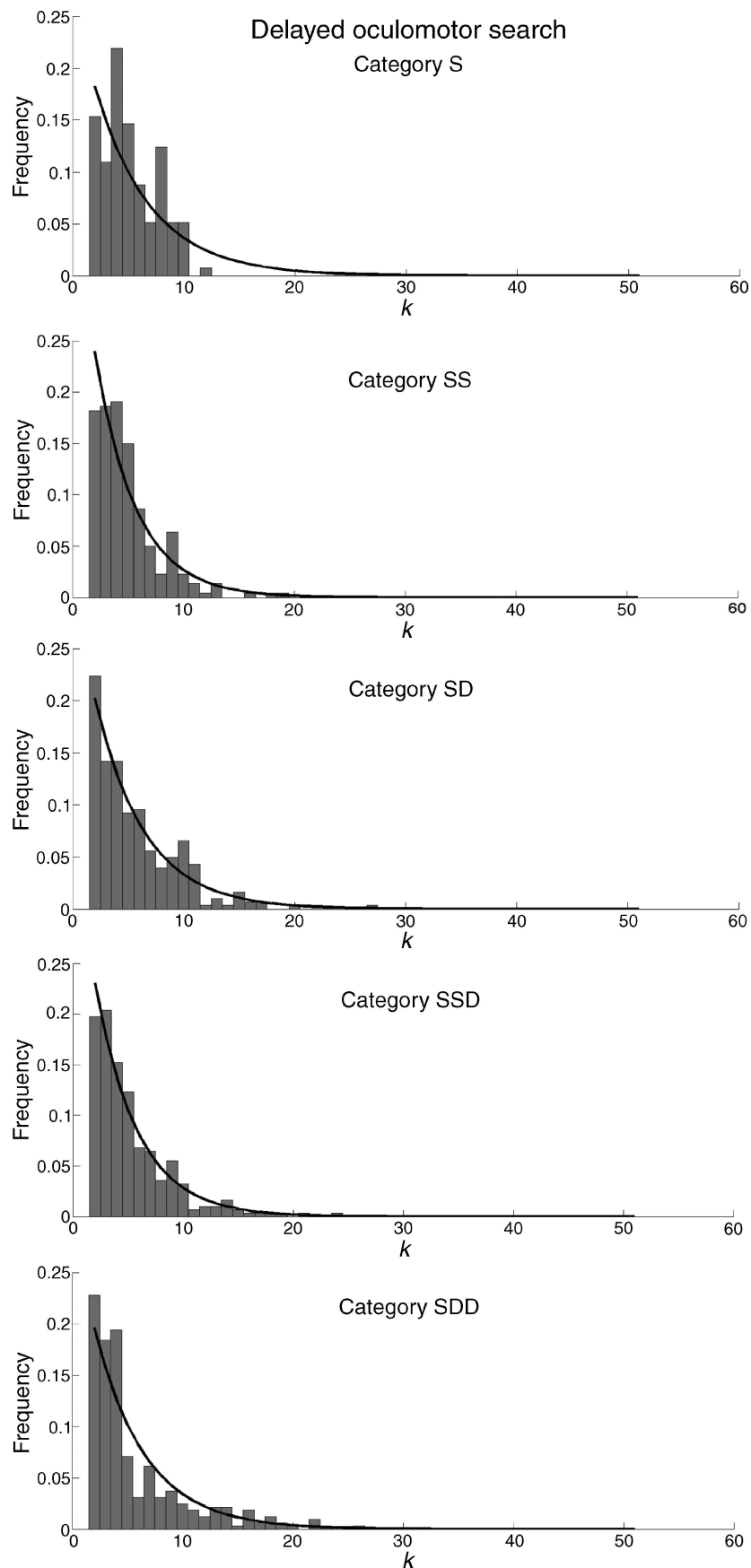


Figure 11. Histograms showing the distributions of  $k$ , the number of locations visited between visiting and revisiting a location, under delayed oculomotor search, Experiment 2. Equation 2 is plotted with the best-fit parameters of  $M$  and  $\epsilon$  for each category (see text for details).

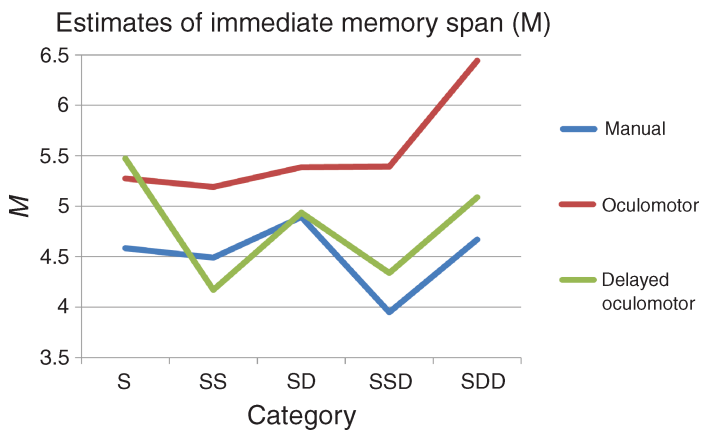


Figure 12. Estimates of  $M$ , the measure of immediate memory span, for each condition and category (see text for details on how  $M$  was obtained). Adding a feature to the category rule (e.g. objects must share a feature) resulted in a decrease in  $M$ , while adding a disjunction (e.g. objects must differ on a feature) resulted in an increase in  $M$ .

made between a revisit of the same object. Following the procedure of Epelboim & Suppes, consecutive visits to the same location were counted as a single visit (consecutive revisits made up <2% of visits in manual search, 8% in oculomotor search with no delay, and 4% in delayed oculomotor search). This meant that  $k \geq 2$ . Parameters  $M$  and  $\varepsilon$  were allowed to vary and were estimated using a minimization search (Matlab function *fmincon*). Values of  $M$  and  $\varepsilon$  were chosen such that the theoretical distribution of  $k$  produced by Equation (2) was as close as possible to the actual distribution of  $k$  in the data. Fit was evaluated using the  $\chi^2$  goodness-of-fit test (see Table 2).

Figure 9 shows distributions of  $k$  for each category rule in Experiment 1 (manual search), along with the theoretical function obtained from fitting the model (Equation 2). Figures 10 and 11 show the same analyses performed for the oculomotor search with and without delays in Experiment 2. Fits of the model to the data were very good in all conditions (see Table 2 for  $\chi^2$  values). Note that the manual search data diverged slightly from what was predicted by the model in that the frequency of revisits for  $k = 2$  was systematically lower than for  $k = 3$  across categories (see Figure 9). This pattern did not appear in Epelboim and Suppes data, nor in the oculomotor search results from Experiment 2, and indicates that some other process, not captured by the model, had some influence on the decisions.

We also confirmed that  $k$  did not vary over the course of a trial. This was done by computing the correlations between each observed value of  $k$  and the proportion of the trial that had been completed at the time of the revisit (where the proportion of the trial that was completed was equal to the ordinal number of the visit divided by the total number of visits in the trial). In oculomotor search,

correlations were near zero (oculomotor with no delay:  $r = .016$ ,  $n = 3969$ ,  $p = 0.29$ ; delayed oculomotor search:  $R = -0.0008$ ,  $n = 1412$ ;  $p = 0.97$ ). There was a small but significant positive correlation between  $k$  and proportion of the trial completed in manual search ( $r = 0.19$ ,  $n = 1951$ ,  $p < 0.001$ ). This was due to the infrequent (9.6%) occurrences of large values of  $k$  ( $k > 10$ ), which were limited to the latter portions of trials.

Values of  $M$ , estimates of the effective span of working memory, are shown in the first column of Table 2 and in Figure 12. Estimates of  $M$  were greater in the no-delay condition (mean = 5.53) than in the delay condition (mean = 4.79; paired  $t = 2.72$ ,  $p = 0.05$ , two-tailed; Table 2). There was no significant difference between  $M$  in the manual and delayed oculomotor search (paired  $t = -1.41$ ,  $p = \text{n.s.}$ ), with both showing significantly smaller estimates of  $M$  than for the oculomotor search without imposed delays ( $F(2) = 6.03$ ,  $p = 0.02$ ). These estimated values of  $M$  are similar to those obtained by Epelboim and Suppes (2001), as well as to the reported average number of items viewed between revisits to the same item by Jacob and Hochstein (2009).

Figure 12 shows that the estimated values of  $M$  varied across the category rules. Before considering the possible implications of these variations in  $M$ , we tested whether the variations were statistically reliable. To do this the model was fit with  $M$  constrained to take on the same value across category rules. The model in which  $M$  was unconstrained fit significantly better than the model in which  $M$  was constrained in both the manual and delayed oculomotor search, verifying the reliability of the effects of category rule on  $M$  (Manual:  $\chi^2 10.595$ ,  $p < 0.001$ ; Oculomotor with delay:  $\chi^2 12.737$ ,  $p < 0.001$ ). The fit of the unconstrained model for the oculomotor condition with no delay was also better, but only marginally so ( $\chi^2 2.334$ ,  $p = 0.08$ ).

Examining the pattern of variation in the values of  $M$  over the category rules (Figure 12) showed that when a feature was added to the rule (Category S vs. SS, and Category SD vs. SSD),  $M$  decreased. On the other hand, when a disjunction was added to the category rule while the number of features remained the same (Category SS vs. SD and Category SSD vs. SDD),  $M$  increased. These variations in  $M$  were as great as about  $\pm 1$  object for manual and delayed oculomotor search; variations were much smaller for oculomotor search with no-delay. The possible implications of these variations in  $M$  will be considered in the Discussion.

## Discussion

The performance of active tasks depends on a series of decisions about when to rely on exploration and when to rely on memory in order to retrieve information about the



content of displays. The category search task was devised to study the role of both cognitive and motor demands in controlling these decisions. The category search task required finding three objects from a set of nine hidden, multi-featured objects that satisfied a given category rule. Cognitive demands of the task were varied by changing the complexity of the category rule, specifically, by adding features or by adding disjunctions to the rule. Motor demands of the task were varied by changing the effector (more effortful arm movements vs. less effortful saccades), as well as by adding time constraints (delay vs. no-delay, in the case of oculomotor search).

Analysis of key aspects of the search patterns (the number of object locations visited, the number of object locations re-visited, and the number of visits between a revisit of the same object location) showed that both the cognitive and motor demands affected the tradeoff between exploration and memory. As the category rules were made more complex by either adding features or adding disjunctions, searchers responded by relying more on exploration, that is, they took more visits and revisits to find the set of objects that satisfied the rule. As motor demands decreased, where the relevant aspect of motor demands was the time required to view an object, rather than the effector (arm or eye), searchers also relied more on exploration, using more visits and a higher rate of visits in the case of oculomotor search with no delay.

These systematic patterns show that decisions about how to tradeoff the use of exploration and memory are not made arbitrarily, but rather are driven by strategic use of resources. We consider possible strategies of resource management below.

### **Strategies during category search: Effects of delay**

One factor driving the strategies during category search was the interest in avoiding the cost of delays. Search in both the manual task and the delayed-oculomotor task, which proceeded at about the same rate (number of objects inspected/second), required about the same number of visits and revisits to find the set, despite the use of different effectors. On the other hand, oculomotor search with no delay resulted in a faster rate of search and more visits and revisits. Adding the delay (750 ms) between fixation on a location and the appearance of the contents resulted in a slower rate of search, that is, fewer objects searched/second, even when the delay duration itself was taken into account.

Delays also influenced the estimates of the span of immediate memory ( $M$ ) in that values of  $M$  were similar for manual and delayed oculomotor search, but larger for oculomotor search with no delay.

One way to interpret the effects of delay is to view the time it takes to perform the action as a resource that needs to be managed. Searchers may have reduced the number of visits in situations where visits were time-consuming

because spending time waiting for display contents to appear was an unproductive nuisance. Alternatively, the effect of delay may reflect strategies for managing memory. With longer delays the probability of previously viewed objects being forgotten may increase. As a result searchers operating with delays may have opted to focus their search on a small number of locations, which would be revisited frequently. This strategy reduces the number of features that have to be preserved over a given period of time, in contrast to a strategy of gathering more information and more features by exploring widely, which is what appeared to characterize oculomotor search with no delay.

### **Strategies during category search: Effects of rule complexity**

Performance in all three motor conditions showed a similar pattern of variation in the number of visits and revisits across the category rules. Adding either features to the rule or adding disjunctions resulted in a more visits and revisits required to find the set. Why? Simulations using an ideal searcher, even one with a fixed capacity memory for the contents of locations, showed only minimal effects of the category rule on the number of visits required to find the set. This shows that the human searchers were modulating the way they used their limited memory across the rules. One way of modulating the use of memory would be by selecting which feature or features to encode from each object (e.g., Droll & Hayhoe, 2007). Although we have no direct measurement of features chosen from a given selected object, the analysis of the span of immediate memory ( $M$ ) (Figure 12) provides the basis for some conjectures.

The analysis of the span of immediate memory showed that  $M$  tended to increase when a disjunction was added to the rule (e.g., SS vs. SD; or SSD vs. SDD). Category rules with disjunctions (i.e., each object in the set must have a different value of the feature), introduce a greater level of cognitive complexity and abstraction (e.g., Feldman, 2000; Jacob & Hochstein, 2008). To deal with the disjunctions, searchers may have used the rule to generate specific, perceptually meaningful hypotheses. For example, given the rule SD (share one feature; differ on another), searchers might decide after viewing one or two objects to search for a set made up of a blue square, blue circle and blue triangle (see also Jacob & Hochstein, 2008). With such a specific and limited search target, searchers may have then reduced the number of features sampled from a given object to those few that were relevant to evaluating the current hypothesis. This strategy would require examination of many objects before a revisit, i.e., a larger value of  $M$ , in an attempt to find a set that conformed to the hypothesis. For rules that lacked disjunctions, on the other hand, such as SS (share two features), searchers could test different perceptual

hypotheses in parallel without having to keep track of as many values of the features per hypothesis as for rules with disjunctions. For example, for rule SS they might be able to search for 3 blue squares or 3 blue circles in parallel. As a result of such a specific, targeted, hypothesis-driven search, finding sets containing disjunctions would require more visits over a larger set of objects, and more visits between revisits to the same object, than rules without disjunctions, which is consistent with the results observed.

According to the above interpretations, the human searcher has limited memory, and must re-visit locations to refresh forgotten details, but (unlike the simulated ideal searcher) is not simply adding all available features to memory and testing all combinations until one emerges that satisfies the rule. The human searcher may be driven by the process of ongoing hypothesis formation and testing, sampling and saving information only to the extent needed to support or refute the current hypothesis. The critical decisions during active tasks, such as searching for a category, are not only whether to rely on memory or exploration, but also how to formulate hypotheses, and how to select the locations that would provide the information that is most useful to testing and evaluating the current ongoing hypothesis driving the search.

### Relation to prior studies

The present results can be compared to the prior study of Jacob and Hochstein (2009), who also examined visual search under conditions that imposed high memory load. In their ‘Identity Search Task,’ searchers looked for a pair of identical, multi-element block patterns among distractor block patterns. They too found revisits to previously viewed items. Jacob and Hochstein (2009) reported there were about 4–5 intervening fixations between a fixation on the same object. This is similar to the estimates of memory span (M) of both Epelboim and Suppes (2001) and the present study (Figure 12), as well as in many prior estimates of the size of immediate or working memory (e.g. Luck & Vogel, 1997). It is interesting that such different tasks produce similar estimates of immediate memory span.

Jacob and Hochstein’s (2009) results also showed frequent fixations on the identical pair prior to selecting the pair, and they speculated that these revisits represented a period of early subthreshold recognition prior to a conscious decision that a pair was found. Jacob and Hochstein (2009) did not specify a mechanism (such as a limited memory buffer) to explain why the revisits were needed. Possibilities include that with each refixation, a representation of the selected candidate pair was being built-up by successively adding more characteristics of the pattern. In this sense, the eye movements in their study were driven by the development of hypotheses—i.e., selection of candidate pairs of blocks—a process we have

suggested was also relevant to category search. Given the low cost of motor exploration in Jacob and Hochstein’s task, there was no incentive or need to tax memory capacity by trying to memorize multiple characteristics of the patterns on each fixation.

## Conclusions

Our results show that cognitive demands and motor demands each affected decisions about how to tradeoff exploration and memory in different ways. This means that decisions about where and when to aim the next movement of eye or arm are based on multiple aspects of performance that must be taken into account concurrently, including the costs and tolerance for delays, the management of memory, and the generation and testing of ongoing hypotheses. These findings lead to two broad conclusions about the strategies used to control motor exploration during active visual tasks:

The first is that strategies of managing memory and exploration are adaptive. There have been numerous examples of how efficient motor performance during active visual tasks is adaptive in that it is driven by monitoring of internal states, including motor variability, motor effort, visual contrast sensitivity, or limits on the span of immediate memory (e.g., Araujo et al., 2001; Ballard et al., 1995; Droll & Hayhoe, 2007; Epelboim & Suppes, 2001; Legge, Klitz, & Tjan, 1997; Motter & Belky, 1998; Najemnik & Geisler, 2005, 2009; Trommershäuser, Maloney, & Landy, 2003; Wu, Kwon, & Kowler, 2010). The category search task studied here is more open-ended and “top-down” than most of the active tasks studied in the past in that there are multiple possible routes to solutions, making the decisions more complex. Nevertheless, despite such complexity, performance depended systematically on both the cognitive and motor characteristics of the task, and thus reflects the application of consistent underlying rules and strategies. These results extend the findings that human beings can adopt efficient adaptive strategies for trading off exploration and memory to situations that impose significant and open-ended cognitive demands.

The second conclusion is that strategies of exploration are linked to hypothesis-testing and decision-making. The consideration of the role of decisions is often given less attention in studies of how people plan exploratory movements—particularly eye movements—because the benefits of eye movements are so strongly linked to overcoming the limits of vision in eccentric retina. However, there are examples where eye movements are preferred, or helpful, when visual resolution for display elements is good enough to support task performance without eye movements (Ko, Poletti, & Rucci, 2010; Kowler & Steinman, 1977), which means that the benefits

accompanying the eye movements may lie elsewhere. Theoretical arguments have been made both for (Ballard, Hayhoe, Pook, & Rao, 1997) and against (Viviani, 1990) a role for eye movements in controlling the temporal sequence of decisions during active tasks. We have argued that the effects we found of category complexity on motor exploration, which was reflected in both the total visits/category, as well as in the pattern of revisits, can be attributed not only to management of memory, but also to links between motor exploration and the ongoing generation, testing and revision of concrete perceptual hypotheses derived from the abstract category rule. Future versions of the category search task, with additional experimental manipulations, such as presenting different features at different locations (e.g., Rehder & Hoffman, 2005), or changing features unpredictably (Droll & Hayhoe, 2007), may provide further insights into how the sequences of exploratory movements are linked to the evolving sequences of decisions and problem-solving strategies during active visual tasks.

## Acknowledgments

Supported by NSF DGE 0549115; NIH EY015522. We thank Jacob Feldman, Manish Singh, Cordelia Aitkin, John Wilder, Brian Schnitzer and Chris Mansley for valuable discussions and for comments on earlier drafts of the manuscript.

Commercial relationships: none.

Corresponding author: Melissa M. Kibbe.

Email: kibbe@rucss.rutgers.edu.

Address: Department of Psychology and Center for Cognitive Science, Rutgers University, 152 Frelinghuysen Rd., Piscataway, NJ 08854, USA.

## References

- Aitkin, C. D., & Feldman, J. (2006). Subjective complexity of categories defined over three-valued features. In R. Sun & N. Miyake (Eds.), *Proceedings of the 28th Annual Conference of the Cognitive Science Society* (pp. 961–966).
- Alvarez, G. A., & Cavanagh, P. (2004). The capacity of visual short-term memory is set both by visual information load and by number of objects. *Psychological Science, 15*, 106–111.
- Araujo, C., Kowler, E., & Pavel, M. (2001). Eye movements during visual search: The cost of choosing the optimal path. *Vision Research, 41*, 3613–3625.
- Baddeley, A. D., & Hitch, G. J. (1974). Working memory. In G. A. Bower (Ed.), *Recent advances in learning and motivation* (vol. 8, pp. 47–89). New York: Academic Press.
- Ballard, D. H., Hayhoe, M. M., & Pelz, J. B. (1995). Memory representations in natural tasks. *Journal of Cognitive Neuroscience, 7*, 66–80.
- Ballard, D. H., Hayhoe, M. M., Pook, P. K., & Rao, R. P. N. (1997). Deictic codes for the embodiment of cognition. *Behavioral and Brain Sciences, 20*, 723–767.
- Bays, P. M., Catalao, R. F. G., & Husain, M. (2009). The precision of visual working memory is set by allocation of a shared resource. *Journal of Vision, 9*(10):7, 1–11. <http://www.journalofvision.org/content/9/10/7>, doi:10.1167/9.10.7. [PubMed] [Article]
- Bays, P. M., & Husain, M. (2008). Dynamic shifts of limited working memory resources in human vision. *Science, 321*, 851–854.
- Brady, T. F., Konkle, T., Alvarez, G. A., & Oliva, A. (2009). Compression in visual short-term memory: Using statistical regularities to form more efficient memory representations. *Journal of Experimental Psychology, 138*, 487–502.
- Coëffé, C., & O'Regan, J. K. (1987). Reducing the influence of non-target stimuli on saccade accuracy: Predictability and latency effects. *Vision Research, 27*, 227–240.
- Droll, J., & Hayhoe, M. (2007). Trade-offs between working memory and gaze. *Journal of Experimental Psychology: Human Perception and Performance, 33*, 1352–1365.
- Droll, J. A., Hayhoe, M. H., Triesch, J., & Sullivan, B. T. (2005). Task demands control acquisition and storage of visual information. *Journal of Experimental Psychology: Human Perception and Performance, 31*, 1416–1438.
- Epelboim, J., Steinman, R. M., Kowler, E., Edwards, M., Pizlo, Z., Erkelens, C. J., et al. (1995). The function of visual search and memory in sequential looking tasks. *Vision Research, 35*, 3401–3422.
- Epelboim, J., & Suppes, P. (2001). A model of eye movements and visual working memory during problem solving in geometry. *Vision Research, 41*, 1561–1574.
- Feldman, J. (2000). Minimization of Boolean complexity in human concept learning. *Nature, 407*, 630–633.
- Feldman, J. (2006). An algebra of human concept learning. *Journal of Mathematical Psychology, 50*, 339–368.
- Gersch, T. M., Kowler, E., Schnitzer, B. S., & Doshier, B. A. (2008). Visual memory during pauses between



- successive saccades. *Journal of Vision*, 8(16):15, 1256–1266, <http://www.journalofvision.org/content/8/16/15>, doi:10.1167/8.16.15. [PubMed] [Article]
- Hardiess, G., Gillner, S., & Mallot, H. A. (2008). Head and eye movements and the role of memory limitations in a visual search paradigm. *Journal of Vision*, 8(1):7, 1–13, <http://www.journalofvision.org/content/8/1/7>. doi:10.1167/8.1.7. [PubMed] [Article]
- Hollingworth, A., & Henderson, J. M. (2002). Accurate visual memory for previously attended objects in natural scenes. *Journal of Experimental Psychology: Human Perception and Performance*, 28, 113–136.
- Hooge, I. T., & Erkelens, C. J. (1998). Adjustment of fixation duration in visual search. *Vision Research*, 38, 1295–1302.
- Inamdar, S., & Pomplun, M. (2003). Comparative search reveals the tradeoff between eye movements and working memory use in visual tasks. In R. Alterman & D. Kirsh (Eds.), *Proceedings of the twenty-fifth annual meeting of the cognitive science society* (pp. 599–604). Boston, Massachusetts.
- Jacob, M., & Hochstein, S. (2008). Set recognition as a window into perceptual and cognitive processes. *Perception & Psychophysics*, 70, 1165–1184.
- Jacob, M., & Hochstein, S. (2009). Comparing eye movements to detected vs. undetected target stimuli in an Identity Search Task. *Journal of Vision*, 9(5):20, 1–16, <http://www.journalofvision.org/content/9/5/20>, doi:10.1167/9.5.20. [PubMed] [Article]
- Kibbe, M. M. (2008). The complexity of a category affects working memory capacity in a search task [Abstract]. *Journal of Vision*, 8(6):1171, 1171a, <http://www.journalofvision.org/content/8/6/1171>, doi:10.1167/8.6.1171.
- Kibbe, M. M., Kowler, E., & Feldman, J. (2009). Oculomotor and manual search compared: The role of cognitive complexity and memory load [Abstract]. *Journal of Vision*, 9(8):1209, 1209a, <http://www.journalofvision.org/content/9/8/1209>, doi:10.1167/9.8.1209.
- Ko, H. K., Poletti, M., & Rucci, M. (2010). Microsaccades precisely relocate gaze in a high visual acuity task. *Nature Neuroscience*, 13, 1549–1553.
- Kowler, E., & Steinman, R. M. (1977). The role of small saccades in counting. *Vision Research*, 17, 141–146.
- Legge, G. E., Klitz, T. S., & Tjan, B. S. (1997). Mr. Chips: An ideal observer model of reading. *Psychological Review*, 104, 524–553.
- Luck, S. J., & Vogel, E. K. (1997). The capacity of visual working memory for features and conjunctions. *Nature*, 390, 279–281.
- Melcher, D. (2001). Persistence of visual memory for scenes. *Nature*, 412, 401.
- Melcher, D., & Kowler, E. (2001). Visual scene memory and the guidance of saccadic eye movements. *Vision Research*, 41, 3597–3611.
- Motter, B. C., & Belky, E. J. (1998). The guidance of eye movements during active visual search. *Vision Research*, 38, 1805–1815.
- Najemnik, J., & Geisler, W. (2005). Optimal eye movement strategies in visual search. *Nature*, 434, 387–391.
- Najemnik, J., & Geisler, W. (2009). Simple summation rule for optimal fixation selection in visual search. *Vision Research*, 49, 1286–1294.
- Oh, S., & Kim, M. (2004). The role of spatial working memory in visual search efficiency. *Psychonomic Bulletin & Review*, 11, 275–281.
- Olson, I. R., & Jiang, Y. (2002). Is visual short-term memory object based? Rejection of the “strong-object” hypothesis. *Perception & Psychophysics*, 64, 1055–1067.
- O’Regan, J. K. (1992). Solving the “real” mysteries of visual perception: The world as outside memory. *Canadian Journal of Psychology*, 46, 461–488.
- Pertsov, Y., Avidan, G., & Zohary, E. (2009). Accumulation of visual information across multiple fixations. *Journal of Vision*, 9(10):2, 1–12, <http://www.journalofvision.org/content/9/10/2>, doi:10.1167/9.10.2. [PubMed] [Article]
- Pothos, E. M., & Chater, N. (2002). A simplicity principle in unsupervised human categorization. *Cognitive Science*, 26, 303–343.
- Pothos, E. M., & Close, J. (2008). One or two dimensions in spontaneous classification: A simplicity approach. *Cognition*, 107, 581–602.
- Rehder, B., & Hoffman, A. B. (2005). Eyetracking and selective attention in category learning. *Cognitive Psychology*, 51, 1–41.
- Shepard, R. N., Hovland, C. I., & Jenkins, H. M. (1961). Learning and memorization of classifications. *Psychological Monographs*, 75, No. 517.
- Sperling, G., & Doshier, B. A. (1986). Strategy and optimization in human information processing. *Handbook of Human Perception and Performance* (vol. 1, pp. 1–65). New York: Wiley-Interscience.
- Trommershäuser, J., Maloney, L. T., & Landy, M. S. (2003). Statistical decision theory and trade-offs in the control of motor response. *Spatial Vision*, 16, 255–275.
- Viviani, P. (1990). Eye movements in visual search: Cognitive, perceptual and motor control aspects. In E. Kowler (Ed.), *Eye movements and their role in visual and cognitive processes (Reviews of Oculomotor Research)* (vol. 4, pp. 353–393). Amsterdam: Elsevier.



- Woodman, G. F., & Luck, S. J. (2004). Visual search is slowed when visuospatial working memory is occupied. *Psychonomic Bulletin & Review*, *11*, 269–274.
- Woodman, G. F., Vogel, E. K., & Luck, S. J. (2001). Visual search remains efficient when visual working memory is full. *Psychological Science*, *12*, 219–224.
- Wu, C. C., Kwon, O., & Kowler, E. (2010). Fitts's Law and speed/accuracy trade-offs during the sequences of saccades: Implications for strategies of saccadic planning. *Vision Research*, *50*, 2142–2157.
- Zeithamova, D., & Maddox, W. T. (2006). Dual-task interference in perceptual category learning. *Memory and Cognition*, *34*, 387–398.
- Zelinsky, G. J., Loschky, L. C., & Dickinson, C. A. (2010). Do object refixations during scene viewing indicate rehearsal in visual working memory? *Memory & Cognition*, published online, doi:10.3758/s13421-010-0048-x.