Spatial Memory And Saccadic Targeting In A Natural Task

M. Pilar Aivar
Department of Psychology, University of Oviedo
Oviedo, Spain

Mary M. Hayhoe
Center for Visual Science, University of Rochester
Rochester, NY, USA

Christopher L. Chizk
Center for Visual Science, University of Rochester
Rochester, NY, USA

Ryan E. B. Mruczek
Center for Visual Science, University of Rochester
Rochester, NY, USA

Previous work on transsaccadic memory and change blindness suggests that only a small part of the information in the visual scene is retained following a change in eye position. However some visual representation across different fixation positions seems necessary to guide body movements. In order to understand what information is retained across gaze positions it seems necessary to consider the functional demands of vision in ordinary behavior. We therefore examined eye and hand movements in a naturalistic task, where subjects copied a toy model in a virtual environment. Saccadic targeting performance was examined to see if subjects took advantage of regularities in the environment. During the first trials the spatial arrangement of the pieces used to copy the model was kept stable. In subsequent trials this arrangement was changed randomly every time the subject looked away. Results showed that about 20% of saccades went either directly to the location of the next component to be copied, or to its old location before the change. There was also a significant increase in the total number of fixations required to locate a piece after a change, which could be accounted for by the corrective movements required after fixating the (incorrect) old location. These results support the idea that a detailed representation of the spatial structure of the environment is typically retained across fixations and used to guide eye movements.

Keywords: Spatial Memory, Saccades, Natural tasks, Eye-hand coordination.

Introduction

In the context of natural behavior, the retinal image is constantly changing because of movements of the eye, head and trunk. As a consequence of these movements, the visual information in different fixations must be coordinated spatially, and information must be preserved in time, to ensure coordinated behavior. However, the nature of the visual information preserved across different gaze positions is still poorly understood. A large body of work on change blindness suggests that very little information is retained from prior fixations. These experiments have shown that observers are very insensitive to changes in the visual scene made during a saccade or other transient, although the same changes are clearly visible if they happen during a fixation on the scene (e.g. Rensink, 2002; Simons, 2000; Simons & Levin, 1997). It is generally agreed that, following a change in gaze position, observers retain in memory only a small number of items, consistent with the capacity limits of visual working memory, together with information about scene “gist”, and other higher level semantic information (Irwin & Andrews, 1996; see review by Hollingworth & Henderson, 2002).

However, in order to understand just what information is retained across gaze positions, it seems necessary to consider the functional demands of vision in ordinary behavior. Although some studies have examined change blindness in the real world (Levin & Simons, 1997; Simons, 1996; Simons & Levin, 1998), most paradigms do not consider how integration across fixations might be needed for natural vision. The importance of task requirements in determining what information is selected and retained in memory has been demonstrated by Triescz et al. (2003). Visual function in experiments that require inspection of images, or simple geometric displays is likely to be fundamentally different from active participation in a real scene, because of different task demands of controlling movements and because the stimulus context is different. For example, viewing a picture of a scene is very different from acting within that scene, simply because the observer needs different information. Another difference that is likely to be important is the nature of the stimulus array. Investigations of change blindness typically involve viewing either two-dimensional pictorial representations of scenes or simple arrays of letters or geometric figures. These displays differ
from normal scenes in their spatial structure. One difference is spatial scale. The visual angle subtended by an image of a room in a typical experimental display, for example, is very different from being in a real room, and it is not clear how such infidelities in spatial scale might affect observers’ representations of the spatial structure of the scene. Depth information introduces an additional level of spatial complexity in normal vision and poses a greater challenge for the visuo-motor apparatus. Moreover, as Xu and Nakayama (2003) have shown, it may also be relevant in determining the capacity of visual short-term memory.

Control of movements is a natural candidate for needing visual representations integrated across fixations. Eye, head, and hand all need to act with respect to a common coordinate system and remain synchronized in time across multiple actions. The reduction in temporal and spatial uncertainty afforded by the continuous presence of stimuli in ordinary behavior allows for the use of visual information acquired in fixations prior to the current one, to plan both eye and hand movements. Chun and Nakayama (2000) hypothesized that implicit memory structures may be needed for guiding attention and eye movements around a scene. They argue that such guidance requires continuity of visual representations across different fixation positions. Such mechanisms do not require conscious intervention, and typically exhibit greater memory capacity, longer durability, and greater discriminability than explicit short-term visual memory (or working memory). This is very different from the memory structures usually hypothesized to span fixations, which are commonly believed to be spatially imprecise (see eg Henderson & Hollingworth, 2003b; Hollingworth & Henderson, 2002; Irwin & Andrews, 1996; Pollatsek & Rayner, 1992). Change blindness studies may therefore underestimate the extent of integration across saccades because the demands of controlling movements are not addressed.

There is evidence to suggest that subjects do in fact build an implicit memory representation of the spatial structure of the display. Chun and colleagues (Chun, 2000; Chun & Jiang, 1998) have shown that subjects are sensitive to the redundancy in visual stimuli and can implicitly learn some aspects of the spatial structure of a scene. In what has been called the “Contextual Cueing” phenomenon, they have shown that visual search is facilitated by prior exposure to the same visual context, as long as the context is informative about the location of the target. This benefit represents a form of implicit learning because subjects could not discriminate old from new contexts in a forced-choice explicit recognition test (Chun & Jiang, 1998). A second phenomenon, called “Priming of Pop-Out”, is another implicit memory mechanism that could be relevant in guiding attention and eye movements. Reaction time to find a target based on its location or features decreases with the repetition of those properties (Maljkovic & Nakayama, 1994, 1996 and 2000), which suggests that the repetition of a property of the target object allows observers to more quickly focus attention on the target. This priming effect occurs automatically and unconsciously, increases with more repetitions of the target property and passively decays over a period of seconds or minutes. It also seems to be important for the efficiency of the saccadic system: as McPeek, Maljkovic and Nakayama (1999) have shown, saccadic latency also decreases when target properties are repeated over trials.

The experiments on contextual cueing and priming of pop-out measured search times with standard experimental displays containing geometric figures. Other evidence suggests that the implicit memory demonstrated in these experiments may also be used in natural environments. Epelboim et al. (1995) found that repeated tapping of a pre-determined sequence of lights on a table led to fewer fixations and faster hand movements with each repetition. This demonstration of learning on a time-scale of minutes strongly implicates the existence of shorter-term visual representations that are built up over fixations and used to guide movements in ongoing behavior. In addition, Hayhoe, Shrivastavah, Mruczek and Pelz (2003) showed that natural eye and hand coordination patterns, when subjects made sandwiches, indicated a need for some representation of the spatial structure of the scene that is built up over different fixations and maintained over a period of a few seconds. They postulated that this representation of the spatial structure of the scene may be important for planning sequences of coordinated movements of the eyes and hands.

The goal of the current investigation was to explore the hypothesis of Chun and Nakayama (2000) and Hayhoe et al. (2003) that in natural vision, precise information about the spatial structure of scenes is retained across gaze position and used in programming movements. In order to mimic the demands of vision in the natural world, while maintaining some degree of experimental control, we used a 3D virtual environment in which observers could pick up and move objects. Observers performed a model copying task, and were required to locate model components, pick them up, and place them in a copy that matched the model. This task contains many of the elements of everyday visually guided behavior, where observers interact with objects in a continuously present scene. The particular question was how saccades are targeted when observers look at a piece in order to pick it up and move it. Do observers use memory of the locations of the pieces from prior views in order to compute the saccade target, or do they locate pieces on the basis of visual search for particular stimulus properties? It is known that observers can make accurate saccades to targets on the basis of memory of stimulus locations when they are required to do so (Colby, Duhamel & Goldberg, 1995; Gnadt & Andersen, 1988; Hayhoe, Lachter & Moeller, 1992; Miller, 1980). However, it is not known whether subjects typically choose
this strategy in natural vision, when the target may be present in the peripheral retina. Indeed, some evidence suggests that visual search functions in a memory-less fashion (i.e. Wolfe, 1999). If, however, observers commonly take advantage of memory from prior fixations in saccade targeting, it seems likely that some representation of the spatial structure of a scene is necessary in addition to memory for objects, scene gist, and other semantic aspects of scenes.

Methods

The goal of this experiment was to test whether there is some cumulative representation of the global context in a scene by looking for facilitation in eye movement targeting by repeating the same spatial pattern over different trials. We also looked for disruption in performance when the configuration was changed. To do this, subjects were asked to copy a model using a set of toy construction pieces (called Baufix). The position of these pieces was kept stable during the first part of the experiment, but changed during the second part. If subjects typically extract a representation of the spatial properties of the environment, the repetition of the positions of the pieces would allow them to generate such a representation. The introduction of changes in the position of pieces should disrupt their behavior in some way, most likely by making the pieces harder to locate when needed. This should be reflected in saccade targeting of the pieces.

Task

A virtual environment was designed with three different areas: the Model area in the central and upper part, the Resource area on the right and the Workspace on the left (see upper part of Figure 1). Wooden parts were simulated to serve as the main elements for the task. Participants were instructed to make copies of a model, which was composed of nine pieces. Eleven additional pieces were placed in the Resource area for the participants to use to complete the task (see lower part of Figure 1). Only nine of those pieces were needed to copy the model, so after finishing the copy there were two pieces left. Previous experiments with a similar task (Ballard, Hayhoe & Pelz, 1995; Ballard, Hayhoe, Pook & Rao, 1997) have shown that participants usually develop a quite stable pattern of eye movements between the areas, like the Model-Pick-Model-Drop pattern described by Ballard et al. (1995, 1997). These stereotyped action patterns allow us to predict subjects’ behavior and manipulate the environment at critical points. In order to copy the model, participants had to make eye, head and hand movements between the different areas: to the model to check its properties, to the resource area to pick up new pieces, and to the workspace to assemble them correctly. The focus of this experiment was on the saccades made to the Resource area for picking up pieces. The location of pieces in the resource area was kept stable in the first part of the experiment, but was altered in the second part every time subjects made an eye movement from the resource to another area. If subjects typically use remembered locations to guide their eye movements, randomly varying the location of pieces should interfere with performance, even if it is not consciously noticed.

Figure 1. Baufix environment. The upper part of the figure shows a general view of the environment. The Model is on the top, the Resource Area is on the right and the Workspace is on the left. The bottom part of the figure shows a close up of the model (left) and the location of pieces in the resource area (right). At the beginning of every trial the locations of pieces in Resource were as shown in the figure. The same model was used for all subjects and trials.

Apparatus

The visual display was delivered via a Virtual Research V8 head mounted display made up of a pair of 1.3 cm LCD panels, each with a resolution of 640x480 pixels. The stereo image was generated by a Silicon Graphics Onyx II with four 250 MHz processors and two Infinite Reality 2 graphics boards, and was updated at 60 Hz. Head position was monitored at 120 Hz with a Polhemus Fastrak 6 degrees of freedom position tracking system, and used to update the display with a latency of less than 50 msec. An ASL Series 501 infrared video eye tracker working at 60 Hz. was integrated into the optics of the helmet and used to monitor position of the left eye. Its accuracy was of about 1 degree. Views of the helmet
and eye-tracker are shown in Figure 2. The ASL signal was recorded and transferred to the SGI, and in addition, a 30 Hz video record of the display was recorded and eye position was superimposed on it. An image of the observer’s eye provided by the ASL was overlaid on the video scene record containing the location of gaze.

Two crosshairs on the eye image indicated the tracker’s calculation of center of the pupil and corneal reflections. When either of these signals was lost, the corresponding crosshair disappeared. This provides a mechanism for checking the scene video for transient track losses and blinks. The movement of the eye can also be seen in the eye image superimposed on the video, providing an additional source of information for identifying fixations and measuring their duration (see Movie 1). In addition, direction of gaze was computed on line and used to change the display in some trials, contingent on gaze.

The virtual pieces were picked up and moved using a second Fastrack sensor held between the thumb and forefinger of the right hand. The position of the hand was visible in the visual environment as a gray cube (see Movie 1). Objects were highlighted in red when the sensor came in contact with them, and the observer picked the objects up or dropped them by pressing the space bar of a keyboard with the other hand.

Movie 1. Example of video record of the display. Gaze position in the scene is indicated with a white crosshair. Hand position is visible as a gray cube. Pieces are highlighted in red when contacted. The eye image provided by the ASL eye-tracker is superimposed in the left upper part of the scene. The two crosshairs on the eye image indicate the eye-tracker calculations of the center of the pupil and the corneal reflection. During the analysis (see below) this sequence of eye movements was classified as Complex Search, because several pieces were fixated in the Resource area before pick up.

Environment properties

The horizontal field of view of the HMD is 54 deg. To ensure the same conditions for all participants, a fixed starting position, 30 cm away from the back wall of the virtual environment was set. At that initial position the Model, Resource and Workspace regions subtended about 18 degrees. The pieces had different colors (orange, yellow, green, red, blue and purple) and three different shapes: there were long bars (3), cubes (2) and bolts (6) (see lower right part of Figure 1). Long bars were approximately 4 cm long and 0.6 cm wide, cubes sides were 0.75 cm and bolts had a diameter of 0.5 cm. At the starting position the long bars subtended about 7.5º along their longest dimension, the cubes about 1.5º and the bolts about 1º of visual angle. The same pieces were used to make a model, which was similar for all groups and trials (see lower left part of Figure 1). Every area was 9 x 9 x 9 cm, except for group 3. In this case

Figure 2. Views of the Virtual Research V8 helmet (top) and the ASL Series 501 eye-tracker integrated into it (bottom).
areas were 18 x 18 x 18 cm and pieces doubled their size, but were placed further away from the subject so as to maintain the same visual angle. Subjects were able to freely move their eyes, head and hand around as they desired.

Subjects

A total of 18 subjects voluntarily participated in the experiment. They all gave their informed consent and were paid for their time. All reported normal or corrected-to-normal vision, were right handed and were naive as to the purpose of the study. There were 12 males and 6 females, and their ages varied between 19 and 24 years. They were randomly assigned to three different groups of 6 subjects each. The research followed the World Medical Association Declaration of Helsinki and was approved by the University of Rochester Research Subjects Review Board.

Experimental Design

Participants were asked to copy the same model several times. The entire set of events involved in copying a particular model is referred to as a trial. In the first part of the experiment (No Swap trials), the different pieces in the resource area were kept in the same position during the whole trial. Their spatial configuration can be seen in Figure 1. In the second part of the experiment (five last trials) the resource pieces started the trial occupying the same positions, but their locations were randomly rearranged every time subjects made an eye movement to a different area of the environment (Swap trials). The rearrangement affected all pieces except the three long bars, which were not moved so that subjects’ attention would not be drawn to the changes. To make the display changes, direction of gaze was computed on line and a random rearrangement of the pieces was triggered 25 frames (417 msec.) after the point where the resource area was outside the field of view of the helmet. Rearrangements of pieces changed the location of every piece in the resource area (except the long bars), but did not affect the spatial configuration of the whole area. That is, although each piece was moved to a new position, it was not possible for a piece to appear in a position that was empty before the rearrangement. The manipulation was done this way to avoid subjects noticing the new positions of the pieces. As Simons (1996) has shown, when a change affects the global structure of the scene it can be easily detected. Although the rearrangements were set to happen in each saccadic movement out of the resource area, in some cases subjects moved too fast out and back into the resource area and the conditions for the rearrangement were not met. For this reason the rearrangements of pieces occurred in about 80% of the pickups during the Swap trials.

The basic design of the experiment consisted of five trials in the No Swap condition followed by 5 trials in the Swap condition. The first five trials, in which the position of pieces was kept stable, gave subjects the opportunity to learn the spatial configuration of the resource area, and also allowed performance to stabilize. In the second part of the experiment the changes were introduced. The model was visible in all trials. Six subjects participated in this version of the experiment (Group 1).

Previous experiments with a copying task have shown that subjects often need to inspect the model for information about the pieces. Ballard et al. (1995); Ballard et al. (1997), and Hayhoe (2000) found that the pattern Model-Pick-Model-Drop was the most frequently used by the subjects to guide their eye movements while finding and moving the different pieces. Hayhoe, Bensinger and Ballard (1998) also reported that model fixations increased duration after introducing changes in model pieces during subjects’ eye movements. To analyze whether participants depended on the model to accomplish the task, a second version of the experiment was tested. It was similar to the basic design (five trials in the No Swap condition followed by 5 trials in the Swap condition) but the model was visible only during the first five trials. Six different subjects participated in this version of the experiment (Group 2).

Following data collection for these two groups, it was observed that performance had not stabilized after the first five trials. The total number of fixations in the model and resource areas decreased steadily over the first five trials and did not appear to have reached asymptote. This raised the possibility that subjects were still learning the spatial configuration of the display. In order to allow more time for subjects’ performance to stabilize, a third version of the experiment was designed, with ten trials in the No Swap condition followed by 5 trials in the Swap condition. As in the basic design, the model was visible in all trials. Six different subjects participated in this version of the experiment (Group 3).

Procedure

Participants received written instructions describing the structure of the environment and their task (to copy the model), but neither the changes in the position of the pieces nor the disappearance of the model were mentioned. No instructions were given as to how to make the copy, so participants were able to organize their actions as they pleased. There was no time pressure to finish the task so participants worked at their own rhythm.

The experimental session started with the calibration of the eye. A calibration grid of nine points subtending about 40 deg. was displayed and participants were asked to look at them consecutively. After calibration was achieved subjects received additional oral instructions about the task and had a few practice trials until they got used to the environment and felt comfortable manipulating the 6D sensor that served as the hand. In those trials, participants saw the same virtual environment but moved
a different set of pieces and did not have a model to be copied. In most cases subjects reported feeling comfortable performing the task after just one practice trial. Then, they received 10 or 15 experimental trials depending on the group they were part of.

Every trial started with all pieces at their original locations in the resource area (as shown in the lower part of Figure 1) and finished when the subject reported that he or she was satisfied with the copy made in the workspace. Calibration of the eye was always checked between trials and the eye was re-calibrated when needed. The manipulations of the display made on the last 5 trials were introduced without warning. All the trials were done in succession in one experimental session, but subjects were free to interrupt the session whenever they felt tired, although always after completing a trial. The eye tracker was recalibrated after each break. Once all trials for one subject were run, a questionnaire was given to analyze subject awareness of the changes introduced during the trials. In the case of group 3 both a recall task and a recognition test were also included to see whether subjects were able to remember the position of the pieces in the resource area. The experimental sessions lasted approximately 45 minutes in the case of groups 1 and 2, and about an hour and 30 minutes when the longer design was used.

Data analysis

**Global patterns of eye movement in the resource area**

When subjects moved their eyes and hand toward the resource area, this area was frequently out of the field of view of the helmet, at least partially. On most of the pickups, subjects landed in the resource area after a large saccade from the workspace or the model area, and then made one or several fixations before picking up a piece. In other cases, their saccades went directly to the piece they picked up, which suggests that subjects were using the remembered location of the piece to guide their saccades. To describe the fixation sequences involved in locating the next piece, the video records were analyzed frame-by-frame. A unique category was used to describe each sequence of eye movements inside the resource area. Each sequence started once the eyes were directed to the resource area from either the workspace or the model areas, and ended once the eyes were moved away from the resource towards one of the other areas.

Six categories were used to describe the different patterns of eye movement while looking for a piece in the resource area (a diagram of most of the categories is shown in Figure 3). Three of them (Direct Movement, Local Search and Complex Search) were main categories that described different ways of locating the next piece in the resource area. Direct Movement (D) was used when the saccade entering the resource area went directly to the next piece to be picked up (see Movie 2). In these cases all fixations happened on the piece being picked up, as there were no other fixations in the resource area. For that reason, this category potentially shows a localization process based on remembered information about the spatial location of the elements in the resource area. (See below, for further discussion.) A second category, Local Search (L), was used when pieces were localized by means of an initial saccade to some empty spot in the resource area and a second small saccade from that point to a specific piece (see Movie 3). This category was included because previous results in a search task (Zelinsky, Rao, Hayhoe & Ballard, 1997) showed that to localize a target subjects made successive saccades in the central area of the scene, each of them moving closer to the target location. Those “center of mass” fixations suggest that, when locating a target, the visual system may first saccade to the approximate location, and then use peripheral information to guide subsequent saccades to the target. The third main category, Complex Search (C), was used to describe patterns with multiple fixations on several different pieces (for an example, see Movie 1). In these cases it was assumed that subjects were using a kind of serial search for the next piece to be moved.

![Figure 3. Diagram of five of the categories used to describe targeting strategies for pickup in Resource.](http://journalofvision.org/4/1/1/)

The introduction of Swap trials produced some special situations that needed to be analyzed, so a category specific for swap trials was added to describe them. Old
to New position (Old) was used when the first fixation in the resource area was made in the old position of a piece before the last swap and the second fixation occurred in the new position of that piece after that swap (occasionally, an additional fixation in another piece happened between this two). This category potentially reveals spatial memory for the positions of the pieces (see below) (see Movie 4).

Movie 2. Example of Direct Movement. Gaze goes directly from the Work area to the yellow bolt in Resource.

Movie 3. Example of Local Search. Both hand and eye move from the Work area to Resource, and stop in an empty spot between the upper and lower rows of pieces before reaching for the green piece in the bottom row.

Movie 4. Example of Old to New movement. At the beginning of the movie the subject is picking up the yellow bolt from the lower row of pieces in the resource area before swap occurs. Note that at this moment the purple bolt is in the upper right position in the resource area. Subject moves to the work area to place the yellow piece and meanwhile swap takes place in the resource area. Subject moves back to resource, and both hand (grey cube) and eye (white crosshair) are directed to the upper right position, where the purple piece was. Now the red cube is there. The eye and hand move to localize the purple piece in the bottom row and pick it up.

Two other categories were used to describe eye movement patterns that happened only occasionally: Next Piece (Next) was used in those cases in which the last piece that was fixated before pick up was picked up on the next visit to the resource; and Other (O), which included any other (infrequent) patterns of search. All unclear or ambiguous movements were eliminated from the analysis (less than 1% of sequences).

Two different researchers (two of the authors) independently categorized all the sequences recorded. Agreement between raters was higher than chance level, as confirmed by Cohen’s kappa (Kappa=0.637, p < .005, N=2076. Kappa shows values between 0 and 1, with 0 meaning that coincidences between raters are at chance levels and 1 that both raters completely agree in their categorization). The proportion of overall agreement between raters was 72.4%. The proportion of agreement specific to each category showed values between 60% (for Local Search) and 89% (for Old to New position). The proportion of agreement for the category Direct Movement was 85%.

After all the sequences were categorized by both raters and their categorizations reviewed and discussed, the proportion of occurrence of the different categories was calculated for every subject and trial, and averaged over the two halves of each experiment. The statistical effect of the introduction of the change was analyzed independently for every category with a repeated measure analysis of variance with two factors: “group” was a between-subjects factor with 3 levels and “swap” was a within-subject factor with 2 levels (no swap vs. swap).
Frequency and duration of fixations in the Resource area

As we have discussed previously, participants had several trials with a stable environment and so were able to learn the spatial location of the pieces needed for the task before changes were introduced. Recent evidence suggests that the introduction of changes in the visual field can affect the frequency and duration of the subsequent fixations (Hayhoe et al., 1998; Henderson & Hollingworth, 2003a; Hollingworth et al., 2001). For that reason it seemed relevant to analyze whether the introduction of swap affected resource fixations or not. Only gaze fixations made inside the resource area were analyzed, and their frequency and duration were calculated. To do so the files recorded during the experiment were analyzed with a program designed in the laboratory to detect saccades and fixations (fixation finder). Based on the data about position of eye and time, this program detected as saccades all data points for which a velocity higher than 60° per second was found, and counted as fixations all groups of consecutive data points under this threshold that reached a total duration of at least 100 msec. An upper velocity threshold of 1000° per second was also set to detect loss of track in the recording. Additional algorithms were included in the program to group consecutive fixations that were less than 2° apart and to estimate lost values. Moreover, the fixations detected for each trial were carefully inspected by hand to verify, and correct if necessary, the program’s analysis, and to add information about the content of each fixation.

After resource fixations were detected, they were also classified in different groups depending on their function in subject’s actions. All those fixations that happened before the next piece to be moved was fixated for the first time were classified in the first group, as fixations locating the piece. The second group, fixations while picking up, included those fixations made on a piece while the hand was moving towards it and while the object was being moved out of the resource area. The third group, other fixations, included the few fixations that could not be classified in one of the previous groups, like those that occasionally happened after locating a piece but before picking up. A last group, fixations after pickup, was introduced to include the occasional fixations that some subjects made on resource pieces while they were already moving the piece to the workspace. These fixations were very infrequent, so they were not analyzed further, but surprisingly appeared only in swap trials.

Because reviewing the tapes to check the fixations found with the program was very time consuming, only three trials were analyzed in this way for each subject: for groups 1 and 2 frequency and duration of resource fixations were analyzed in trials 4, 5 and 6, and for group 3 the same analysis was made in trials 9, 10 and 11. These are the two trials prior to the Swap manipulation, and the first trial after Swapping was introduced. These trials were chosen specifically to analyze the effect of changes in the environment (Swap condition) on eye movements and gaze fixations. A repeated measures analysis of variance with a within-subjects factor (trial, with three levels) and a between-subjects factor (group, with three levels) was used to analyze the effect of the introduced changes in the environment and the differences between groups. An independent analysis of variance was used for the two main kinds of fixations: fixations locating the next piece, and fixations while picking up.

Results

Global patterns of eye movements in Resource area

As described above, a system of categories was used to describe the different patterns of eye movements that the subjects used to localize the pieces in the resource area. To find out how subjects were locating the different pieces and which strategies they preferred, the proportion of occurrence of the different categories over each part of the experiment was calculated. Figure 4 shows the average proportion of occurrence of the different categories in the first five trials of the experiment, before the swap condition was introduced. Each bar shows the values for each of the three groups, averaged over the first five trials (ten in the case of group 3). Note that the conditions for the different groups were essentially identical in the first part of the experiment. Complex Search, where subjects made a series of fixations to locate a piece, occurred in 35-40% of the cases, and both Direct Movement and Local Search occurred in around 25% of the cases. The frequency of the other categories was lower, as can be seen in the rightmost part of Figure 4.

Figure 5 shows changes in the frequency of the four main strategies as a function of trial number, averaged over the three groups for the first and last five trials of the experiment. During No Swap trials, the main change in category use as subjects repeated the task was the reduction of the frequency of complex search. This suggests some general increase in familiarity with the spatial layout in the resource area. In the case of group 3, category use remained stable between trials 5 to 10, and is not shown.

The introduction of changes in the environment (Swap condition) produced significant variations in the occurrence of some of the patterns of eye movements used by the participants to find pieces. As can be seen in Figure 6, the proportion of occurrence of both “Direct” movements and “Complex Search” decreased with the introduction of swapping, but the category that showed a fixation in the old position of the piece (Old to New Position) occurred with a relatively high frequency (around 20% of the sequences). It is important to notice that Old to New movements are only possible during Swap trials,
because they require a fixation in the old position of a piece before a change in position. “Other” movements also increased their frequency during the Swap condition. The differences in occurrence reached statistical significance, as can be seen in Table 1. There was no significant trend in the category frequencies as a function of trial number after swapping was introduced, as can be seen in the right part of Figure 5.

Although the model was removed for subjects in group 2 and the design was longer in the case of group 3, there were no significant differences between groups in the proportion of occurrence of the four main categories, as can be seen on the right column of Table 1.

![Figure 4](http://journalofvision.org/4/1/1/)  
**Figure 4.** Proportion of occurrence of each category over all No Swap trials for each group. Proportions were calculated independently for each subject and then averaged for each group. In the case of Group 3 ten trials contributed to the values showed, instead of five, because a longer design was used. Error bars show the standard error of the mean.

![Figure 5](http://journalofvision.org/4/1/1/)  
**Figure 5.** Average frequencies of the four main categories for the first and the last five trials of the experiment (for group 3 only trials 1 to 5 and 11 to 15 were included in the figure). Averages were calculated over the three groups. Error bars show the standard error between the three groups.

![Figure 6](http://journalofvision.org/4/1/1/)  
**Figure 6.** Changes in the proportion of occurrence of the different categories with the introduction of swap. The plot shows the differences between the proportions calculated over all no swap and over all swap trials. Positive values refer to increases in the occurrence of that category after the introduction of Swap; negative values refer to decreases in occurrence. Error bars show the standard error of the mean. Code: D – Direct movement; L – Local search; C – Complex search; Old – Old to New position; Next – Next piece; O – Other categories.

<table>
<thead>
<tr>
<th></th>
<th>Effect of Swap</th>
<th>Effect of Group</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>F Value</td>
<td>Prob.</td>
</tr>
<tr>
<td>Direct Movement</td>
<td>28.931</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Local Search</td>
<td>2.265</td>
<td>.153</td>
</tr>
<tr>
<td>Complex Search</td>
<td>5.320</td>
<td>.036</td>
</tr>
<tr>
<td>Old to New position</td>
<td>228.473</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Next Piece</td>
<td>1.084</td>
<td>.314</td>
</tr>
<tr>
<td>Other</td>
<td>16.794</td>
<td>.001</td>
</tr>
</tbody>
</table>

**Table 1.** Effects of Swap and Group in the repeated measures ANOVAs calculated for each of the categories. The analysis was run on the proportion of occurrence of each category over all no swap and swap trials for each subject. None of the interactions reached statistical significance.

**Interpretation of Category Use**

Complex Search was the most commonly occurring category of search pattern. On these occasions it seems unlikely that subjects are using spatial memory to target a specific piece, but rather go to some random location in the resource and then search for a particular piece on the basis of visual features (see Movie 1). In the Local Search category, subjects could be using either visually based
search, or memory. In the latter case they may target a remembered location, and then correct the movement (see Movie 3). However, for a significant proportion of the pickups during no swap trials (around 25%), subjects landed directly on the piece to be picked up after a single large saccade into the resource area (Direct Movement), and it is tempting to assume the saccade was programmed on the basis of spatial memory. We examined the video records in each of these movements to find if the landing point was in fact visible in the peripheral retina at the point of initiation of the saccade. We found that in about 49% of the movements the target was not visible, so these are clearly programmed on the basis of spatial memory (see Movie 2). In the remaining cases, the target was visible, so current retinal image information, or some combination of memory and current visual information, may have been used to program the movement. It is also possible that subjects were not targeting a piece of a particular color, and simply happened to pick up the piece they landed on by chance.

In general this seems unlikely, as subjects copied the patterns in a very reproducible order, suggesting that each visit to the resource area was for the purpose of locating a piece of a particular color and shape. Subjects’ tendency to construct the copy in the same order is shown in Figure 7. The proportion of times each piece was placed at a specific moment over trials is plotted against order of placement. As can be seen in the plot, average performance in our experiments was very close to always placing the pieces in the same order. The calculation of Goodman and Kruskal’s Gamma as a measure of ordinal association showed that in almost 80% of the cases it was possible to predict correctly which piece would be placed at that specific moment based on previous performance (Gamma=.798, p < .001).

The interpretation of the Old to New category search pattern, when the resource pieces moved around in the swapping condition, is similarly ambiguous. As in the case of the Direct movements during no swap trials, a large proportion of these saccades were made when the initial landing point was out of the visual field. Inspection of the video records revealed this was the case for 77% of the Old to New movements. Note that this is a somewhat higher proportion than for the direct movements. These movements must have been programmed on the basis of information from prior views. Of the remaining Old to New movements, when the landing point was visible, we cannot definitively conclude that the saccade was programmed on the basis of memory information. However, it was often observed that the hand movement targeted the same location, and arrived at about the same time as the eye, and made the same corrective movement to the new piece, which strengthens the suggestion that a particular spatial location was being targeted (see Movie 4). Note that if this interpretation is correct, the saccade target selection process must use memory information exclusively, even though inconsistent visual information is present in the retinal image because a different piece is now at the old location.

![Figure 7](http://journalofvision.org/4/1/1/)  
**Figure 7.** The proportion of times each piece was placed at a specific moment over trials is plotted against order of placement (each position in the sequence has a different color). If the same order were used in each trial the plot would show values of one at the diagonal, and ceros everywhere else. As we can see, average performance in our experiments was very close to always placing the pieces in the same order.

**Frequency and duration of resource fixations**

As described above, fixations in the resource area, on trials just before and just after the Swapping manipulation was introduced, were also analyzed to see whether the introduction of changes in swap trials had any effect in their frequencies or durations.  

**Functional differences between resource fixations**

An analysis of the fixations in the resource area showed that actions while finding the next piece to be moved could be organized primarily in two different phases: first subjects search the area until they localize the next piece to be moved, and then they fixate it while controlling the hand to pick it up. Both fixations while locating next piece and while picking it up are the most frequent, and together account for 85% of the fixations made on the resource area. Other fixations (for example, fixations on other pieces after location but before pickup) only appeared in 13.6% of the cases and their frequency differed between subjects. Because of the higher between-subject variability, these fixations were not analyzed further.

As can be seen in Table 2, the mean frequencies of fixations locating and picking up a piece are quite similar, with a mean value of around 12 fixations per trial; that is, a little more than one fixation, on average, of
each kind for every piece that was moved. However, fixation durations are very different depending on their function; fixations made while picking up have a duration of 2000 msec on average, while fixations made while locating the next piece are shorter (around 300 msec on average). This result presumably reflects the different requirements of locating the piece and guiding the pickup.

<table>
<thead>
<tr>
<th>Group</th>
<th>Trial</th>
<th>Locating piece</th>
<th>Picking piece</th>
<th>Other fixations</th>
</tr>
</thead>
<tbody>
<tr>
<td>G 1</td>
<td>T4</td>
<td>12.5</td>
<td>338.5</td>
<td>12.5</td>
</tr>
<tr>
<td></td>
<td>T5</td>
<td>12.16</td>
<td>297.4</td>
<td>11.83</td>
</tr>
<tr>
<td></td>
<td>T6</td>
<td>14</td>
<td>358.8</td>
<td>12.16</td>
</tr>
<tr>
<td>G 2</td>
<td>T4</td>
<td>12.33</td>
<td>275</td>
<td>12.5</td>
</tr>
<tr>
<td></td>
<td>T5</td>
<td>16</td>
<td>293.6</td>
<td>12.5</td>
</tr>
<tr>
<td></td>
<td>T6</td>
<td>19.33</td>
<td>285.4</td>
<td>12.83</td>
</tr>
<tr>
<td>G 3</td>
<td>T4</td>
<td>9.16</td>
<td>444</td>
<td>11</td>
</tr>
<tr>
<td></td>
<td>T5</td>
<td>8</td>
<td>382.1</td>
<td>10</td>
</tr>
<tr>
<td></td>
<td>T6</td>
<td>10.83</td>
<td>326.5</td>
<td>10.83</td>
</tr>
</tbody>
</table>

Table 2. Mean frequency and mean duration (in msec.) of fixations in resource area per trial, classified depending on their function (locating next piece, picking it up, or other). Mean values were first calculated separately for each subject and then averaged for each group.

Epelboim et al. (1995) also segregated fixations into search (or locating) fixations and those guiding the tapping movement, with a similar difference in durations. Given that there is no haptic feedback, and contact with the piece is signaled by a color change, it is to be expected that the process of picking up the piece in the virtual environment should be quite long. Dependence of fixation duration on the specific visual information required has also been described in other experiments using natural tasks (Hayhoe, Shrivastava, Mruczek & Pelz, 2003; Land, Mennie & Rusted, 1999).

**Effect of the introduction of changes in the environment**

During Swap trials, multiple changes were introduced in the resource area every time participants made a saccade away from it. If saccades into the resource area preparatory to picking up a piece are planned using spatial information from prior fixations in the resource area, then such a manipulation should affect those saccades made after a swap. To analyze this effect, frequencies and durations of fixations in the resource area were obtained and compared between the last two trials before swap and the first swap trial. Only fixations made while locating and while picking up were analyzed. Long bars were excluded when analyzing locating fixations because they did not vary their position in swap trials.

As can be seen in Figure 8, the introduction of changes in the position of the pieces in the Resource area produced an increase in the number of fixations made while locating the next piece, but not while picking it up. This increase in the frequency of locating fixations with the introduction of swap reached significance (see Table 3). Within subjects contrasts showed that the first trial after swap was significantly different from the two trials prior to swapping (trial 4 vs. trial 6: p=.005; trial 5 vs. trial 6: p=.002), while the last two trials before swap did not differ (trial 4 vs. trial 5: p=1). Although the increase in the number of locating fixations was higher for group 2, there were no significant differences between the groups in the increase in frequency after swap (F2 =2.042, p=.164; tested by a one way ANOVA on the effect of group over the differences in frequency between the last trial without swap and the first one with swap).

![Figure 8. Mean frequency of fixations per piece moved. Two kinds of fixations are shown; while locating (left) and while picking (right) the next piece. The last two trials in the No Swap condition are compared with the first trial in the Swap condition. Different lines show the data corresponding to the three different groups of subjects in the experiment. Error bars show the standard error of the mean.](image-url)
Thus the introduction of changes in the environment significantly increased the number of fixations made by the subjects while looking for the next piece to be picked up, but did not affect fixations made while controlling the hand movements needed to pick up that piece, or once the next piece had been located. This means that when the pieces were in a stable location in the resource area, subjects took advantage of that to aid search. Interestingly, the introduction of changes only affected the frequency of the fixations made during the locating phase, but not their duration (see Table 3).

**Effect of practice**

As can be seen in Figure 8, all three groups of subjects showed the same effect with the introduction of changes during swap trials. In all cases, after the introduction of changes in the environment the frequency of fixations while locating pieces increased. However, as can be seen in Table 3, there were also significant differences between groups in the frequency of fixations while picking up the next piece. Specifically, subjects of group 3, who experienced 10 trials in the task before swap was introduced, showed significantly fewer fixations while picking the next piece than the other two groups (Post-hoc analysis using Tukey contrast showed significant differences between group 3 and groups 1 (p=.015) and 2 (p=.002)). There were no significant interactions between the effects of group and swapping. This result suggests that practice decreased the number of fixations during pickup. However, there were no differences between groups in the mean duration of fixations, which suggest that practice does not reduce the amount of time needed to process the contents of each fixation.

**Origin of extra fixations in the resource area**

After finding that the introduction of changes in the environment (swap condition) produced an increase in the number of fixations made in the resource area while locating the next piece, we were interested in analyzing how such an effect was related to the different patterns of eye movements that subjects could use. To do so, all fixations made while locating the next piece were classified based on the category that was used to describe the sequence they were in. This analysis showed that the new fixations in the resource area that appeared during swap trials resulted from the use of the "Old to New" strategy to find the next piece. Figure 9 plots the same data as in Figure 8, for frequency of fixations made while locating a piece (black line), together with that value with the fixations resulting from the Old to New category removed (red line). Thus it can be seen that all the extra fixations resulting from introducing swaps came from those sequences in which subjects went to the old, remembered position of the piece, and then to the new one after the change. This result strongly supports the interpretation of the Old to New fixations as being based on memory. A paired samples t test also confirmed a significant difference in the mean number of fixations between both cases, when the fixations resulting from the Old to New category were or were not included in the total (T_{17}=4.873, p<0.001).

![Figure 9](http://journalofvision.org/4/1/1/)

Figure 9. Frequency of fixations per piece while locating it, including (black line) and excluding (red line) those fixations that occurred in Old to New sequences. Error bars show the standard error of the mean.

**Effect of Model disappearance**

As discussed above, previous experiments using a model copying task have shown that subjects often rely on multiple fixations in the model to both pick up and drop a piece from the resource area (Ballard, Hayhoe & Pelz, 1995; Ballard, Hayhoe, Pook & Rao, 1997; Hayhoe, 2000; Hayhoe, Besinger & Ballard, 1998; Karn & Hayhoe, 2000). This behavior suggested that subjects prefer a minimal memory strategy, where they fixate the model in preference to using visual memory of its properties.

The present experiment differed from those experiments, in that the same model pattern was repeated in each trial, thus giving subjects longer exposure to the same patterns. We measured the frequency with which subjects looked at the model for each trial of the experiment (but without calculating how many fixations took place in each look). This analysis showed that over the trials of the experiment subjects of all groups needed to look less and less often to the model (see Figure 10). This decrease appeared in both possible cases, when looking at the model before picking up a piece (that is, prior to locating it), and when looking at it after a pick up (in the way to the work space). At the beginning of the experiment, subjects needed to look at the model around
16 times (taking together before and after pickup looks) in the course of copying it: that is almost twice per piece. After 5 trials, subjects looked at the model about 6 times, or less than one time per piece. After 10 trials, they were looking at the model about one time per two pieces. 

copied the model. This suggests that the piece they landed on was explicitly targeted, as opposed to a situation where observers simply picked up the piece that they accidentally landed on. Second, about half the movements were initiated when the landing point was outside the field of view. Such movements must rely on spatial memory for target selection. For the rest of the movements, the target was usually close to the edge of the visual field (about 25 deg eccentricity) at the point when the saccade was initiated, usually as a consequence of an ongoing head rotation towards the resource area. In these cases it is possible that the visual features of the object played a role in target selection. A stronger argument for the use of spatial memory comes from the Old to New category of gaze movements in the latter part of the experiment when the pieces changed locations every time the observers looked away (see Movie 4). The fact that observers do not pick up the piece they land on, but instead move immediately to fixate and pick up the piece that had previously been in that location, strongly suggests the use of spatial memory in targeting\(^7\). This strategy accounted for about 20 percent of the gaze movements into the resource area during swap trials. A final argument implicating spatial memory is the increase in fixations required to locate a piece when the pieces changed location following each pickup. About 2 to 5 additional fixations were made in the resource area in the course of the trial when the pieces were changing locations (that is, about 0.3-0.8 extra fixations per piece, see Figure 8). Importantly, all these additional fixations came from those sequences in which observers landed on the old location and then had to find the piece in the new location.

One of the goals of this experiment was to evaluate the demands that natural behavior places on memory from prior fixations. Although it seems fairly clear that observers used memory from prior fixations to target movements back to the resource area, it was by no means the most common strategy. Only about 25% of movements in the first 5 trials were Direct movements to the pieces, and only about 20% of movements went to the Old location during the subsequent Swapping trials. On about half the movements, observers made multiple fixations around the resource area before locating a piece for pickup (Complex Search). Thus, while observers commonly use spatial memory to program the movements, they are more likely to make a large saccade to the general region, and then to search for the piece in a local region, presumably using visual features. On the other hand, on those trials when memory-based targeting is implicated (Direct and Old to New movements), the spatial precision is quite impressive, since the gaze changes were 20 to 30 deg in magnitude, and the targeting precision approximately 2-3 deg. Land et al (1999) have also noted excellent accuracy in very large gaze changes to regions outside the field of view (greater than 90 deg), that land within a few degrees of the target. The need to ori-

Discussion

The current experiment provides evidence that memory of the spatial structure of a scene is retained across gaze positions and is used in saccadic targeting in the course of natural behavior. When selecting a target for gaze changes into the resource area, for the purpose of picking up a piece needed for copying, observers frequently made saccades that fell directly on the piece they then picked up (see Movie 2). This occurred on about 25% of the gaze changes. Although these gaze changes cannot be definitively identified as memory guided, there are several arguments for this interpretation. First, subjects were quite consistent in the order with which they
ent to regions outside the field of view in natural vision (for example, moving around within a room) provides a rationale for storing information about spatial layout. The fact that many of the Direct and Old to New saccades were actually to regions currently visible in the retinal image suggests that spatial memory information is not used exclusively for locations outside the field of view. Consistent with this, Edelman, Cherkasova and Nakayama (2002) and Kristjánsson and Nakayama (2003) have observed that subjects are able to locate and saccade to targets that are un-resolvable in the peripheral retina, provided they have been fixated previously. This suggests that spatial memory aids target selection for objects within the field of view.

An advantage of a strategy that uses memory information, whether or not the target is within the field of view, is that it may minimize the number of movements (and time) required to locate a piece. Although minimizing the time to locate a piece might not always be particularly critical, another, and possibly more important advantage is that it allows early planning of head and hand movements. We have found that, in this experiment, observers initiate the head movement to the resource area on average about 400 msec before the eye movement. The head movement is initiated about 200 msec before the eye (Hayhoe, Aivar, Gaines & Jovancovic, 2003). Typically, in response to a visually presented target, head and hand movement initiation lag behind the eye by 100 msec or more (Abrams, Meyer & Kornblum, 1990). The early initiation of the movements has the consequence that the head and hand both arrive close in time to the arrival of the eye in the resource area. Thus a significant role for memory for the spatial layout of a scene is probably for early planning and coordination of the eye, head, and hand movements.

The time course of the memory for spatial structure observed in this experiment is difficult to evaluate. The reduction of the frequency of Complex Search patterns over the first 5 trials is consistent with observers building up long-term memory representation of the layout of the resource over a period of tens of minutes and of the order of a hundred fixations. A similar reduction in the number of fixations required to locate items for tapping was observed by Epelboim et al. (1995). However, the reduction in Complex Search frequency was not accompanied by an increase of the Direct movements, that explicitly implicate memory use. The frequency of the Direct strategy remained fairly stable across the first 5 trials (see Figure 5). The decrease in frequency of Complex Search patterns might therefore reflect other aspects of learning, such as learning what piece to select. Another possibility is that subjects may learn general aspects of the spatial structure, such as the location of the resource relative to the workspace, the general layout of the resource, and potential locations for pieces, rather than the location of specific pieces. This information may aid search without necessarily contributing to the frequency of Direct or Old to New movements. Alternatively, the reduction of Complex Search frequency might reflect subjects’ use of some amalgam of current visual information with memory signals, that facilitates search, but does not clearly implicate spatial memory as the Direct and Old to New strategies do.

Although the overall reduction in number of fixations as a function of trial number in both model and resource areas points towards some kind of accrual in long term memory, the memory revealed by the occurrence of Old to New fixations may be of shorter duration. Since all the pieces change position every time the subject looks away from the resource area, it is only possible to use location information from the immediately prior visit to the resource area. If observers depended only on long-term memory, one might expect to see no evidence for memory-based targeting at all. However, the Old to New strategy accounted for an extra 2-5 fixations per trial, and about 20% of the movements. This suggests that subjects base their targeting on memory from the immediately prior visit to the resource area. Several fixations, and about 5 secs intervene before the return to the resource, so the information is not rapidly decaying, but nonetheless it does not appear to reflect accumulation over the entire experiment: when a recognition task was presented at the end of the experiment to the subjects in group 3 none of them were able to select the picture showing the position of the pieces at the beginning of the trial. In the tea making task, Land et al. (1999) also noted a number of instances where objects were found more easily when they had been fixated a few seconds previously. The current observations provide further evidence that memory across fixations is needed as a basis for motor planning and coordination.

It is also a little surprising that the frequency of the Old to New strategy remains fairly constant across the 5 trials where the pieces changed every trial (see Figure 5). Subjects do not appear to reject or move away from the memory-based strategy over time, despite the incidence of landing on the wrong piece. It is also interesting to note that the incidence of the Complex Search strategy does not appear to increase (in fact it decreases) when the pieces change locations in the Swap trials. This suggests that the memory responsible for the overall reduction in Complex Search as a function of trial number is not the same as that responsible for targeting the pieces in the Direct Search and Old to New movements. That is, the memory used for targeting the resource pieces for the most part is short-term, though there may be some undetermined facilitation from long-term accrual. In this respect the time course is more comparable to priming of pop-out, which is effective over a few trials (Maljkovic & Nakayama, 1994, 1996 and 2000), than to contextual cueing, which accrues over blocks of trial (Chun & Jiang, 1998).

Although the use of memory to guide saccadic targeting in this experiment is not the dominant strategy, it is
clearly a significant aspect of performance. Movements to the resource area are frequently planned using visual information acquired several seconds previously during prior fixations in the region. This means that memory representations integrated across saccades must include precise spatial information that can be used for saccade planning, in addition to scene gist and a small number of object files, as previously proposed (e.g., Irwin 1991; Irwin & Andrews, 1996; Irwin, Zacks & Brown, 1990; O’Regan, 1992; O’Regan & Levy-Schoen, 1983). Other evidence shows that information about the spatial organization of scenes is preserved across fixations. For example, De Graef and Verfaille show encoding of spatial relationships of “bystander” objects that are not the target of a saccade (de Graef et al., 2001; Verfaille et al., 2001). Hayhoe et al. (1992) also showed integration of very precise spatial information across saccades that served as a basis for spatial judgements. O’Regan (1992) and Irwin (1991) have postulated that there is some integrated representation of the scene, but suggest that the representation of spatial information is imprecise and that the representation is semantic in nature. The evidence presented here, however, supports the suggestion of Chun and Nakayama (2000) that the spatial information cannot be imprecise, but must be able to support high precision movements. Other evidence also suggests that the original proposals of O’Regan (1992) and Irwin and Andrews (1996) probably underestimate the extent of the memory across saccades. For example, Hollingworth and Henderson (2002), Irwin and Zelinsky (2002), Melcher (2001), Melcher and Kowler (2001), and Tatler, Gilchrist and Rusted (2003) have all demonstrated robust visual memory representations of multiple objects and their locations in images of complex scenes. However, it is difficult to determine the precision of the spatial information retained in these experiments because a partial report technique was used to explore memory of the objects in the scene. In our experiments, the accuracy of both eye and hand movements in the Direct and Old to New strategies suggests that the spatial information retained is quite precise. However, more research is needed to more directly assess the precision of the spatial information retained in memory.

In summary, despite compelling evidence from change blindness studies, our results suggest that there are implicit memory mechanisms implied in saccadic guidance and movement control that can retain precise spatial information about the objects in the scene. Such mechanisms are useful for the adequate coordination of eye and hand movements, and so need to be studied under complex paradigms that try to get closer to the conditions of vision and action in natural behavior. Thus using simple tasks to measure trans-saccadic memory does not reveal the extent to which memory is required for ordinary behavior.

Acknowledgments

This research was supported by National Institutes of Health Grants EY 05729 and RR 06853. M.P. Aivar was supported by a doctoral FPU grant from the Ministry of Education and Culture of Spain (AP97-10903352), and by a research grant from the University of Oviedo. Portions of this article are based on a dissertation submitted by M.P. Aivar in fulfillment of the requirements for the PhD degree at the University of Oviedo.

We would like to thank Brian Sullivan and Diane Kucharczyk for their assistance with data collection and analysis; Tomás R. Fernández and Jose Carlos Sánchez for helpful discussions on these issues; and all subjects that participated in this experiment for their collaboration.

Commercial relationships: none.
Corresponding author: M.P. Aivar.
Present address: Erasmus Medical Center. Dr Molewaterpl 50. 3015GE - Rotterdam, The Netherlands. Email: p.aivar@erasmusmc.nl

Footnotes

1 While doing the task subjects habitually maintained only one or two of the areas inside the visual field of view. To be able to see the whole configuration subjects needed to move their heads backwards so that the three areas will fit the visual field.

2 Note that there is a small proportion of Direct fixations during the Swap condition. These fixations could result from subjects picking up a long bar (long bars did not change position during swap trials), or could also appear in those sequences during swap trials in which swap did not occur (although it was set to happen every time, swaps only occurred on 80% of the sequences). It is also possible that subjects may have opted to pick up the new piece that occupied the position they landed on after a Direct saccade.

References


