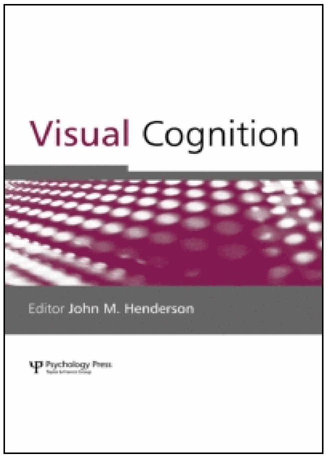


Visual Cognition
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t713683696

# Objects or locations in vision for action? Evidence from the MILO task

Todd S. Horowitz [a]; Ian M. Thornton [b]
[a] Brigham and Women's Hospital, Harvard Medical School, Boston, MA, USA
[b] University of Wales, Swansea, UK

First Published on: 01 March 2007
To cite this Article: Horowitz, Todd S. and Thornton, Ian M. (2007) 'Objects or locations in vision for action? Evidence from the MILO task', Visual Cognition, 16:4, 486 - 513
To link to this article: DOI: 10.1080/13506280601087356
URL: http://dx.doi.org/10.1080/13506280601087356

VISUAL COGNITION, 2008, 16 (4), 486–513

Psychology Press
Taylor & Francis Group



# Objects or locations in vision for action? Evidence from the MILO task

Todd S. Horowitz
*Brigham and Women's Hospital, Harvard Medical School, Boston, MA, USA*

Ian M. Thornton
*University of Wales, Swansea, UK*

In the Multi-Item Localization task (MILO; Thornton & Horowitz, 2004), observers are asked to find an ordered sequence of targets. We can measure the influence of both past actions and future plans on search for the current target. Our previous work with static search arrays found evidence for both retrospective and prospective memory. Responding to a target eliminated its influence on subsequent responses, while observers consistently planned ahead at least one item into the future. Here we asked whether these effects were based in location- or object-based reference frames. We used dynamic arrays in which observers had to search for multiple moving targets. Our results suggest that observers can still plan ahead effectively in this dynamic environment, indicating that future target objects can be tracked as they change position. However, memory for previous targets is essentially eliminated, suggesting that locations, not objects, were being tagged in our previous work.

We recently introduced a new task for exploring the spatiotemporal context of search (Thornton & Horowitz, 2004). In the Multi-Item Localization (MILO) Task, observers are presented with a stimulus array containing multiple target and distractor items (typically alphanumeric characters), and are asked to click on an ordered sequence of targets. The MILO task thus differs from the classic visual search paradigm in two important ways. First,

Please address all correspondence to Todd S. Horowitz, Visual Attention Laboratory, Brigham & Women's Hospital, 64 Sidney Street, Suite 170, Cambridge, MA, USA. E-mail: toddh@search.bwh.harvard.edu

We would like to thank Randall S. Birnkrant, Sarah B. Klieger, Andries Hof, and Jessica Wilner-Reid for assistance with data collection and analysis. TSH was supported by NIH grant MH65576, and IMT by the Max Planck Institute for Biological Cybernetics, Tübingen, Germany. We would also like to thank Steve Franconeri and two anonymous reviewers for helpful comments on the manuscript.


http://www.psypress.com/viscog DOI: 10.1080/13506280601087356

requiring observers to interact with targets, as opposed to a categorical detection or identification response, more closely mimics our natural interactions with the environment. Second, the use of a sequence of targets, as opposed to a single target, makes it possible to explore the influence of both past and future actions within the space of a single trial. Figure 1 illustrates a typical MILO trial. The observer is first presented with a cue display, indicating the sequence of targets. Once the observer signals that he/she has memorized the sequence, a search array is presented, consisting of a set of ovals marked with target and distractor characters. The primary dependent variable is the Sequential Reaction Time (SRT), the time to click on target $n$ minus the time to click on target $n-1$.

Two simple manipulations allow us to measure the influence of past and future actions independently. To explore retrospective memory, we compare a condition in which targets vanish from the screen as soon as the observer responds to them to a condition in which the targets remain on the screen. Physically removing the targets from the screen simulates a perfect retrospective memory. To the extent that performance is inferior when the old targets remain on the screen, we infer that old targets still have some influence on behaviour.

Prospective memory can be revealed by altering the target sequence mid-trial. For example, in Thornton and Horowitz (2004), we introduced a manipulation in which each response triggered the shuffling of future target

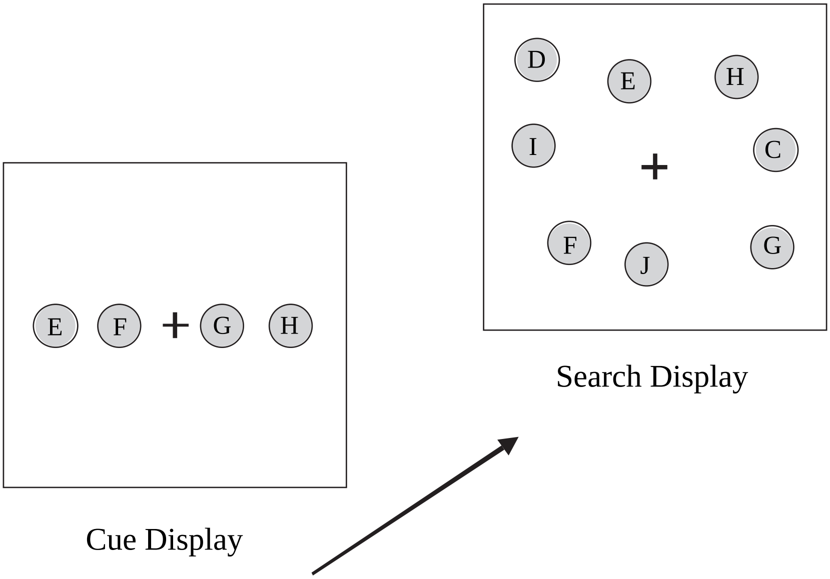


**Figure 1.** Example stimulus displays from the MILO task.

and distractor items. The physical layout of these ovals on the screen was not altered, with the shuffle manipulation simply randomizing which item was in which oval. A key feature of this manipulation was that old targets were left unchanged. This allowed us to keep prospective and retrospective effects separate. The shuffle manipulation prevented observers from gaining any advantage from planning ahead. Aspects of the data that changed when future target locations were shuffled in this fashion could thus be attributed to prospective memory.

Both retrospective and prospective effects can be seen in Figure 2, which is redrawn for the data reported in Thornton and Horowitz (2004). This graph shows SRT for a sequence of four targets within a field of four distractors. In Panel A, SRT decreases with each response, a pattern that would be predicted in the case where targets vanish after a response (open symbols), as the physical set size is being reduced. The fact that a similar pattern is obtained when old targets remain on the screen (filled symbols) illustrates the efficiency of the retrospective memory. The large drop in SRT between items 1 and 2 in both curves illustrates the prospective effect. This gap can be eliminated by shuffling the position of future target and distractor items (Panel B). Under these conditions all subsequent SRTs remain at the level of the initial SRT.

All of the experiments reported by Thornton and Horowitz (2004) used static search displays. That is, the physical location of the target and distractor ovals never varied during the course of a trial (though the contents of the ovals were sometimes shuffled, as noted above). The purpose of the

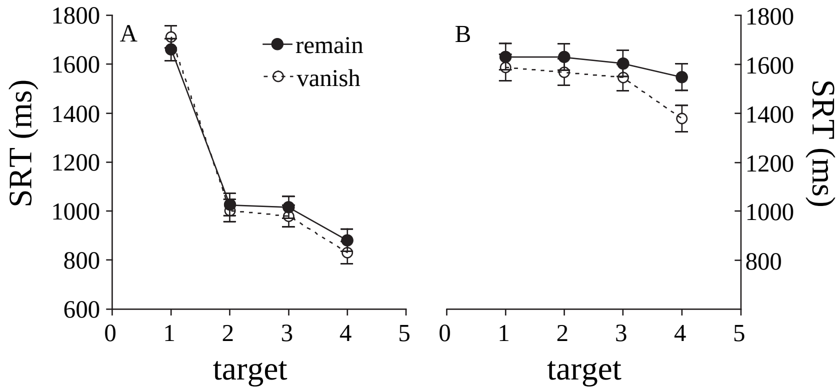


**Figure 2.** Results from the MILO paradigm using static displays (redrawn from Thornton & Horowitz, 2004). SRTs are plotted as a function of target number and condition. Open symbols denote the "vanish" condition, filled symbols the "remain" condition (see text). Panel A depicts the results when stimulus positions remained constant throughout a trial. Panel B depicts the results when we shuffle the locations of distractors and targets that have not yet been responded to.

current work was to examine MILO performance in situations where search displays were dynamic rather than static, with all ovals constantly in motion.

Our primary purpose in introducing motion to the MILO task was to study the frame of reference in which retrospective and prospective effects take place. Our previous data (Thornton & Horowitz, 2004) are compatible with both location-based and object-based reference frames. That is, observers might have used memory for target locations, or some form of location-invariant representation. With the addition of motion, the ability to use location-based coding should be eliminated or at least severely disrupted. If similar patterns of retrospective and prospective performance were obtained with dynamic displays, this would argue for a location-invariant representation. Is this a plausible outcome?

A substantial literature, beginning with the seminal work of Duncan (1984), supports the notion that attentional selection can act on objects instead of (or at least in addition to) locations (Abrams & Law, 2000; Davis, Welch, Holmes, & Shepherd, 2001; Schendel, Robertson, & Treisman, 2001; Watson & Kramer, 1999). A number of selection-related effects, such as negative priming (Shapiro & Loughlin, 1993) and inhibition of return (IOR; McAuliffe, Pratt, & O'Donnell, 2001; Tipper, Jordan, & Weaver, 1999; Tipper, Weaver, Jerreat, & Burak, 1994), have been shown to be coded in both location- and object-centred coordinate systems. Indeed, coding in multiple reference frames seems to be a common representational strategy (Danziger, Kingstone, & Ward, 2001; Ghafouri, Archambault, Adamovich, & Feldman, 2002; Lamy & Tsal, 2000).

We reasoned that both retrospective and prospective memory for targets in the MILO task would likely use multiple coding systems. Search and manipulation of real-world objects would seem to call for an object-centred frame of reference. However, studies demonstrating object-centred effects often use sparse displays with many fewer objects than the MILO paradigm. Object-centred representations might be more computationally expensive, leading to capacity limitations in the number of objects that can be tracked. Our best guess as to what this capacity might be comes from studies of Multiple Object Tracking (MOT; Pylyshyn & Storm, 1988). Such studies have established that the visual system can track several independently moving items simultaneously; the usual capacity estimate is 3–5 items (for a review, see Cavanagh & Alvarez, 2005). However, this varies inversely with item speed; at the low velocities we employed in these experiments, observers should be able to track roughly 7–8 items (Alvarez & Franconeri, 2005).

In Experiment 1, we took the basic MILO paradigm and set the ovals in motion at 1.2 degrees visual angle per second (°/s). The results were broadly similar to those shown in Figure 2A, except that SRTs declined more steeply for the vanish condition than for the remain condition, indicating that retrospective memory was impaired. Experiment 2 replicated this finding, using a larger set size to more clearly expose the impact of motion on retrospective performance. In Experiment 3, we introduced the shuffle manipulation, described above, which is known to disrupt prospective memory in the MILO task. Retrospective memory was again impaired in Experiment 3. More importantly, the shuffle manipulation removed the gap between the first and subsequent items, confirming that observers were still able to plan ahead in Experiments 1–2. In Experiment 4, we slowed the motion to 0.9, 0.7, or 0.5°/s, but there was no sign that retrospective memory became more effective as the speed was reduced. In Experiment 5, we reduced the velocity even further, to 0.25°/s, without changing the results. Finally, in Experiment 6, we introduced rigid motion, so that the configuration of items remained unchanged throughout a trial; this did not alter the results. Together, these results suggest that retrospective memory in the MILO task is location based, while prospective memory is location invariant.

## GENERAL METHODS

All the experiments shared the same basic methodology. The Method sections for individual experiments will specify only deviations from this procedure. Experiments 1 and 2 were conducted at the Max Planck Institute for Biological Cybernetics, Tübingen, Germany, Experiments 3, 4, and 6 at the Visual Attention Laboratory, Brigham and Women's Hospital, Boston and Cambridge, MA, USA, and Experiment 5 at the Psychology Department, University of Wales, Swansea, UK.

### Participants

All observers had normal or corrected-to-normal vision and were naïve as to the purpose of the research until after the experimental session. Observers provided informed consent, either according to procedures approved by the Max Planck Institute, the Partner's Healthcare Corporation Institutional Review Board, or the University of Wales, depending on where the study was conducted. Observers in Tübingen were paid 8 €/hour for participation, observers in Cambridge and Boston $10/hour, and observers in Swansea £8/hour.

### Equipment

Software was custom written using routines based on work by Pelli and Zhang (1991), Rensink (1990), and Steinman and Nawrot (1992). Observers were seated at a standard viewing distance of 60 cm in front of the monitor. All experiments were conducted on Macintosh computers and 21-inch colour monitors with a refresh rate of 75 Hz. Screen resolution was either $1152 \times 780$ (Experiments 1, 2, and 5) or $1024 \times 640$ (Experiments 3, 4, and 6). These resolution differences led to slight variation in the size and speed of stimuli, which will be noted in the appropriate experiments.[1] In the next section, we assume a setting of $1152 \times 780$.

## Stimuli

Stimuli were presented on a uniform, dark grey background. A central viewing area subtending approximately 12.0° was delineated by a black outline square. The search items were light grey oval patches, subtending approximately 1.0° in diameter. The border of the oval was drawn in black, and each item contained a single uppercase letter, also drawn in black. The bounding box of each letter subtended 0.6°. The letters were always present in the ovals, with the following exception: When an observer clicked on any target, the centre of *all* ovals was redrawn in the background grey for 50 ms. This global transient was introduced in order to mask the local transients when letters were shuffled between ovals (see Experiments 3–5 for details). The global transient does not affect any of the experimental manipulations (Thornton & Horowitz, 2004).

In the search display, initial positions of all eight items were randomized within the central viewing area on a trial-by-trial basis. The only constraint on this position was that each item must be at least 1.0° away from all other items and the edge of the viewing area. As soon as the search display appeared, all items began to move in pseudorandom directions at a fixed velocity of 1.2°/s. The animation was achieved by shifting each oval by one pixel after a delay of two screen refresh cycles. Initial path lengths and directions were computed separately for each item. Directions were selected from one of eight cardinal compass points (i.e., N, NE, E, SE, S, SW, W, NW), and path distance was randomly varied within the range of 2.6–3.4°. When an item had traversed its initial path distance, a new direction was randomly selected and a new path length assigned.

When items approached the edge of the central square, the direction of motion was simply reversed, so that the object appeared to bounce away from the edge in a predictable manner. When two items approached each other they did not bounce, but overlapped, with one item passing over and

[1] While these changes were not intentional, we have previously varied the size of displays within this range with no influence on performance. In Experiments 4 and 5, we explicitly manipulated speed, again with no obvious effects over a much larger range.

occluding the other. The occlusion precedence (which item was in front) was not predictable on the basis of target sequence order, as items were randomly assigned to ovals. Importantly, the current target item was always selectable during such occlusion.

## Task

Each trial began with the cue phase, in which the four target items were displayed in the centre of the screen, with two items on either side of a central fixation cross (see Figure 1). The target letters always formed an alphabetic sequence (e.g., E, F, G, H) that was randomly chosen on each trial. The sequence could fall anywhere in the Roman alphabet between the letters C and X. Observers were free to view the target sequence as long as they liked; they then initiated the search phase of the trial by pressing the spacebar.

The search display contained the four target items plus four distractor items. The distractor items flanked the target set in the alphabet (e.g., C, D, I, J). The observer's task was to locate each of the target letters *in sequence*, beginning with the first letter (i.e., E) and ending with the last letter (i.e., H). Observers responded by clicking anywhere within the target oval using a standard computer mouse. A trial was terminated, and the screen blanked, either when the four-item target sequence had been successfully completed or when an error occurred. An error could occur either because a mouse click occurred in the wrong oval ("sequence errors") or because a mouse click occurred in the background (i.e., outside of any oval, "click errors").

## Procedure

There were two main stimulus conditions. In the *vanish* condition, each target item disappeared after the observer clicked on it. In the *remain* condition, targets remained on the screen and continued to move after they were clicked on.

Observers completed two blocks of each condition. Each block contained a minimum of 30 trials. Trials in which an error occurred were replaced, so that a block only finished when observers had successfully completed 30 trials. Each observer thus completed a minimum of 120 experimental trials, taking short breaks after each block. Block order was counterbalanced, with half of the observers completing the two vanish blocks before moving on to the remain blocks and the other half seeing the opposite order.

Observers were first shown several example trials. They were then familiarized with the response method and allowed to practise until they were comfortable using the mouse. The error conditions were illustrated to all observers, and they were informed that a block would only terminate when they had successfully completed 30 trials. Therefore, while speed of response was emphasized, they were also motivated to avoid errors.

Note that while observers were provided with a fixation cross at the start of the trial, the cross disappeared during the search phase. Observers were free to move their eyes, and we did not record eye position.

### Data analysis

Median SRT was the primary dependent variable, defined as the time to click on target $n$ minus the time to click on target $n-1$. Error rates were arcsine transformed before analysis, and we converted the transformed means back into percentages before reporting. We also computed within-subjects confidence intervals (Loftus & Masson, 1994).

## EXPERIMENT 1

The purpose of Experiment 1 was to determine whether the previously reported ability to "ignore" old targets in the MILO task was based on memory for the *location* of an object or for the object itself.

### Methods

Twelve observers from the Tübingen community participated in this study. Screen resolution was set to $1152 \times 780$.

### Results

Errors are shown in the right-hand panel of Figure 3. There were substantially more click errors than sequence errors. However, neither error type varied with condition: click errors, $F(1, 11) = 1.0$, $p > .10$, $MSE = 0.02$; sequence errors, $F(1, 11) < 1.0$, $MSE = 0.01$.

SRTs are plotted in the left-hand panel of Figure 3 as a function of target number. SRTs were an average of 62.9 ms slower in the remain condition than in the vanish condition, $F(1, 11) = 7.3$, $p < .05$, $MSE = 13{,}092$. SRT also declined significantly with target number, $F(3, 33) = 239.2$, $p < .000001$, $MSE = 6719$. There was no interaction, $F(3, 33) < 1.0$, $MSE = 3158$.

A glance at Figure 3 suggests that the main effect of target number was powerfully driven by the drop from the first target to the second target. However, the analysis did not change substantially when we eliminated the initial SRT. The main effect of condition was still significant, $F(1, 11) = 5.4$,

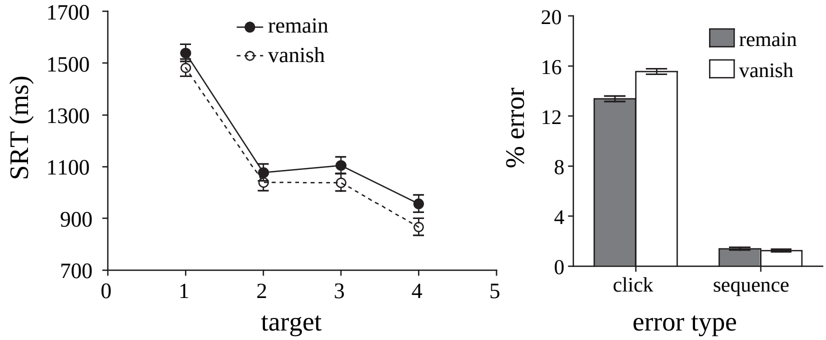


**Figure 3.** Left-hand panel plots SRT medians from Experiment 1 as a function of target number. Filled symbols denote the remain condition, open symbols the vanish condition. Right-hand panel plots percentage error as a function of error type. Grey bars denote the remain condition, open bars the vanish condition. In this and all subsequent figures, error bars indicate within-subjects 95% confidence intervals (Loftus & Masson, 1994).

$p < .01$, $MSE = 14{,}056$, as was the main effect of target number, $F(2, 22) = 40.0$, $p < .00001$, $MSE = 4723$. The interaction term increased, but not to the level of significance, $F(2, 22) = 1.3$, $p > .10$, $MSE = 2988$.

We also computed SRT × Target number slopes (based on the last three SRTs). The slopes for the remain condition averaged –60.5 ms/target, compared to –86.0 ms/target for the vanish condition (CI ± 21.0 ms/target). This difference was marginally significant, $F(1, 11) = 3.6$, $p = .08$.

## Discussion

An initial comparison between Figures 2 and 3 would suggest that the current data quite closely resemble those we collected with static stimuli (Thornton & Horowitz, 2004). Particularly prominent is the large drop in SRT after the first target, a feature that we have previously attributed to prospective memory; during the search for the first target, observers were able to obtain some information about the location of subsequent targets. This ability would appear to be intact, even when the objects are in motion. We will return to this issue in Experiment 3.

In the retrospective domain, however, motion does appear to have had an effect on behaviour. In our previous studies with static stimuli, the vanish and the remain conditions always yielded virtually identical patterns of results, suggesting highly efficient memory for past targets. In the current experiment, SRTs in the remain condition were consistently slower than in the vanish condition. More importantly, the decline in SRT with target order was also generally shallower. Such disruption implies that retrospective performance in the MILO task is sensitive to object location rather than being based exclusively on object identity.

Motion did have a very clear impact on the error rates. Although there was no difference between the vanish and the remain conditions, the overall level of errors was considerably larger than that observed with static arrays (Thornton & Horowitz, 2004). In our experiments with static stimuli, overall errors seldom exceeded 5%, compared to 15–17% in the current experiment. However, the bulk of the errors in this experiment were click errors rather than sequence errors.

A sequence error occurs when a successful motor movement is aimed at the wrong object. Thus, we propose that sequence errors measure cognitive errors, where the observer at some stage selected the wrong object for action. Click errors, on the other hand, represent motor errors in aiming and executing mouse movements. There may be some cases in which conflict at the cognitive level leads to an inaccurate motor movement, or motor errors that by chance land in the wrong object. However, we suspect these cases are rare. Thus, we hold that sequence errors measure cognitive difficulty and click errors motor difficulty.

In this framework, then, it is significant that the bulk of the errors in this experiment were click errors (see the right-hand panel of Figure 3). The rate of sequence errors here (1–2%) is comparable with data from static arrays. However, while click errors in the static arrays were also around 1–2%, here they are 13–16% of trials. We suggest that these errors simply reflect the difficulty of hitting a moving target. In agreement with this idea, subsequent experiments with slower speeds will produce lower click error rates, while sequence errors will remain uniformly low. Importantly, these effects do not interact with the reference frame effects that are the primary focus of this paper.

Overall, the current results suggest that it is the retrospective aspects of performance that are most disrupted by motion. Memory for old items may thus rely on location- rather than object-based representations. With a target sequence of only four items, however, it may be premature to exclude the possibility of object-based tagging, especially since the effect was statistically weak. We designed the next experiment to obtain a clearer picture of the relationship between vanish and remain conditions in the context of motion.

## EXPERIMENT 2

One simple way to gain more information about the relationship between retrospective performance in the vanish and the remain conditions is to increase the number of responses that have to be made. In Experiment 2, we increased the target sequence length from four to six items. If motion does eliminate the ability to retrospectively tag old targets, then SRTs for vanish and remain should clearly diverge as a function of position within the target sequence.

## Method

Fourteen observers from the Tübingen community were paid for participation in this study. We followed the General Methods except for the following two changes. First, the number of target items was increased from four to six. Second, to maintain the ratio of target to distractor items, the number of distractors was also increased to six, giving an overall display set size of twelve. All other aspects of the display and task were identical to Experiment 1.

## Results

Two observers were removed from the analysis, one due to exceptionally slow RTs, and one for very high error rates. Error rates for the remaining 12 observers are shown in the right-hand panel of Figure 4. Neither type of error varied with condition, both $F(1, 11) < 1.0$, $MSE = 0.02$.

SRTs are plotted in the left-hand panel of Figure 4. As in Experiment 1, SRTs were significantly longer for the remain condition than for the vanish condition, $F(1, 11) = 19.1$, $p < .005$, $MSE = 17{,}858$. SRTs also declined with target number, $F(5, 55) = 76.7$, $p < .000001$, $MSE = 14{,}396$, but this decline was steeper for the vanish data, $F(5, 55) = 6.1$, $p < .0005$, $MSE = 6452$.

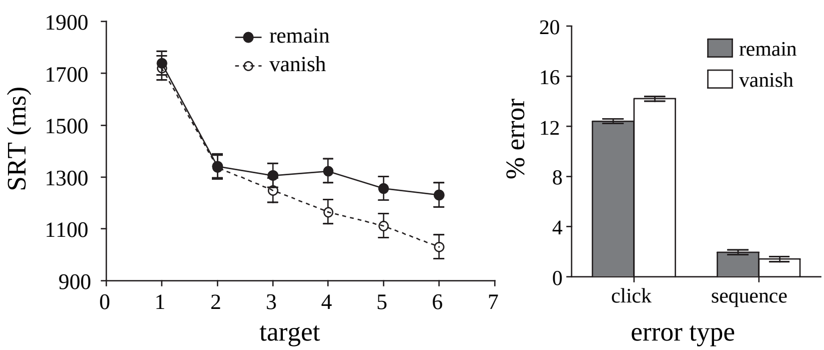


**Figure 4.** Left-hand panel plots SRT data from Experiment 2 as a function of condition and target number. Filled symbols denote the remain condition, open symbols the vanish condition. Right-hand panel plots percentage error as a function of error type. Grey bars denote the remain condition, open bars the vanish condition.

As in the previous experiment, the drop from Target 1 to Target 2 seems to account for much of the effect of target number, so we reanalysed the data, excluding RTs to the first target. The same effects were obtained: main effects of both condition, $F(1, 11) = 17.8$, $p < .005$, $MSE = 21{,}609$, and target number, $F(4, 44) = 10.7$, $p < .000001$, $MSE = 14{,}872$, as well as an interaction, $F(4, 44) = 7.0$, $p < .0005$, $MSE = 5445$.

In order to explore the interaction, we computed the SRT $\times$ Target number slopes for the two conditions (again excluding the initial RT). The slope in the vanish condition, at –75.2 ms/target, was significantly steeper than the –27.4 ms/target slope (CI $\pm$ 18.9 ms/target) in the remain condition, $F(1, 11) = 15.5$, $p < .005$, $MSE = 887$.

## Discussion

With the extended target sequences in Experiment 2, the difference between vanish and remain conditions became much clearer. As expected, the vanish condition SRTs declined sharply as the physical set size was reduced. The corresponding slope for the remain condition was substantially shallower. This finding indicates that retrospective memory was not as efficient for moving targets as for static targets, where the slope of the two functions has previously been shown to be essentially identical (Thornton & Horowitz, 2004). As in Experiment 1, there was a large drop in SRT after the first response for both vanish and remain conditions, suggesting that initial planning was not disrupted by object motion. Error rates were again higher than previously observed with static displays (Thornton & Horowitz, 2004), but did not vary as a function of condition.

# EXPERIMENT 3

Another way in which to focus more clearly on the retrospective aspects of the MILO task is to eliminate the ability to plan ahead. As noted above, the signature of prospective memory in this paradigm is the large decrease in SRT between the first and second targets. This signature is clearly present in the data from the previous two experiments. We hold this pattern to be the consequence of observers picking up advance information about future targets during search for the first targets. Experiments 1 and 2 thus suggest that the prospective aspects of the MILO task are less tied to location and may indeed operate in an object-based manner. We return to this issue in the General Discussion, but the current paper will continue to focus on the retrospective aspects of the MILO task.

With static arrays, the ability to plan ahead was easily disrupted by *shuffling* the contents of future target items at the time of a response. In this way, information about the position of such items was rendered useless. This manipulation removed the large gap between responses 1 and 2, but did not affect retrospective performance (Thornton & Horowitz, 2004, Experiment 2). Here we replicated that manipulation with dynamic arrays. There were two main goals. First, to reaffirm that gap between the first and second responses does relate to forward planning. Second, to provide a picture of retrospective performance in the absence of such forward planning. We predicted that the SRT gap between the first and second items would be substantially reduced, and the difference between vanish and remain conditions would be more clearly seen, as all four responses would now be of the same order of magnitude.

## Method

Sixteen observers were recruited from the Brigham and Women's Hospital Visual Attention Laboratory volunteer panel. The procedure was as described in the General Methods, with the following exceptions. After the observer clicked on a target, the locations of all distractor letters and subsequent target letters were randomly reshuffled among the current stimulus locations. Targets that had already been clicked on were unchanged. It is important to note that the overall layout of the display was not changed, only the contents of the ovals (i.e., the letters) were updated. As previous studies had indicated that the shuffle manipulation is quite demanding for observers, we also returned to using four instead of six items.

Screen resolution was set to $1024 \times 768$ in this experiment. Thus, the overall size of the stimulus area was 13°, and individual targets subtended 1.1°. Items moved at 1.4°/s.

## Results

Errors are shown in the right-hand panel of Figure 5. Neither click nor sequence errors differed by condition, both $Fs(1, 15) < 1.0$, $MSE = 0.01$. The left-hand panel of Figure 5 plots SRTs as a function of condition and target number. As in previous experiments, SRTs in the remain condition were longer than in the vanish condition, $F(1, 15) = 17.7$, $p < .001$, $MSE = 27{,}257$, and there was a decline in SRT with target number, $F(3, 45) = 30.7$, $p < .0001$, $MSE = 5168$, which was steeper for the vanish than the remain data, $F(3, 45) = 18.6$, $p < .0001$, $MSE = 4046$.

Unlike in the first two experiments, we did not observe an elevated RT to the first target. This is a typical result from the shuffle manipulation (Thornton & Horowitz, 2004). Therefore, we did not reanalyse data without the first RT. The SRT × Target number slope in the remain

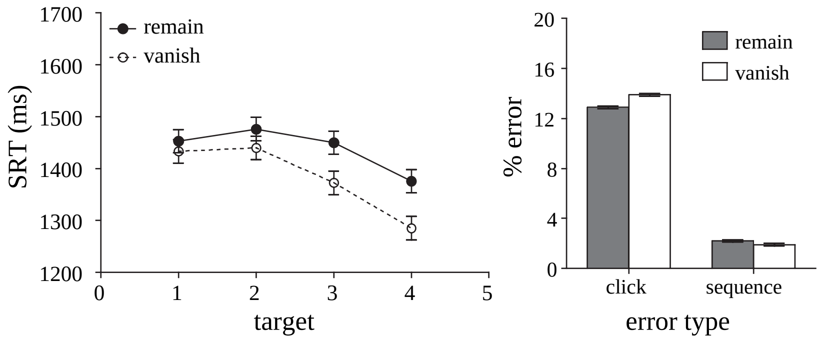


**Figure 5.** The left-hand panel plots median SRT data from Experiment 3 as a function of condition and target number. Solid symbols denote data from the remain condition, open symbols data from the vanish condition. Right-hand panel plots percentage error for the two error types. Grey bars denote the remain condition, open bars the vanish condition.

condition was $-26.0$ ms/target, which was significantly shallower than the $-51.2$ ms/target (CI $\pm$ 11.4 ms/target) slope in the vanish condition, $F(1, 15) = 11.2$, $p < .005$, $MSE = 458$.

## Discussion

Experiment 3 yielded two major findings. First, as we predicted, the large drop in SRT from the first to second targets was abolished, and the SRT $\times$ Target function became quite linear for the vanish condition. This suggests that prospective memory functions similarly in both static and dynamic displays and that, in Experiments 1 and 2, such forward planning used an object-based coordinate system.

Second, we observed very little in the way of retrospective effects in the current experiment. There was a substantial difference between the remain and the vanish conditions, both in terms of mean SRT and the slope of the SRT $\times$ Target function. While the slope in the vanish condition was somewhat shallower than the $-70$ to $-90$ ms/target range seen in the previous two experiments, it was still twice as steep as the remain slope, which was slightly shallower than that observed in Experiment 2. This suggests that observers were largely failing to keep track of old targets. In the absence of forward planning, this pattern was much easier to see.

# EXPERIMENT 4

The difference between the vanish and the remain conditions in Experiments 1–3 suggests that, in the context of motion, observers are largely unable to ignore old targets. This suggests that our previous retrospective findings were based on memory for location rather than memory for objects. In the next two experiments we begin to explore the precision involved in such retrospective memory representations.

The object motion we have used so far had a constant velocity of either 1.2 or 1.4°/s. If we assume an average response time of around 1200 ms, it is clear that previous targets will have moved completely outside of their initial locations by the time a subsequent response is made. Is a complete lack of overlap with initial position necessary in order to abolish the retrospective effects? Furthermore, average SRTs became longer over the first three experiments, and the retrospective effect correspondingly smaller. These two facts might be related: The longer an observer takes to make a response, the further each old target moves from its remembered location.

In Experiment 4, we tested the hypothesis that the strength of retrospective memory would vary inversely with item speed. We measured the strength of retrospective memory as the inverse of the slope of the SRT × Target function in the remain condition. Our prediction is that this slope will be steeper for slower speeds than for faster speeds, because targets will stay closer to their remembered locations.

## Method

Twenty-three observers from the Brigham and Women's Hospital Visual Attention Laboratory volunteer panel participated in this study. Experiment 4 was identical to Experiment 3, except for the following modifications. First, there were two groups. One group of 12 observers ran only the vanish condition, the other group of 11 observers ran only the remain condition. This modification was made to maintain the overall number of trials performed by each observer within the range used in Experiments 1–3. Second, we varied the speed of object motion in separate blocks. Motion was always slower than that used in previous experiments, and was constant at either 0.5°, 0.7°, or 0.9°/s.

## Results

Error data are shown in Figure 6C. Neither type of error rate differed between groups: click errors, $F(1, 21) = 2.1$, $p > .10$, $MSE = 0.07$; sequence errors, $F(1, 21) < 1.0$, $MSE = 0.03$. Click errors increased with speed, $F(2, 42) = 12.8$, $p < .00005$, $MSE = 0.01$, more rapidly so for the observers in the vanish condition than for those in the remain condition, $F(2, 42) = 4.5$, $p < .05$, $MSE = 0.01$. The simple effect of speed was significant for the vanish observers, $F(2, 22) = 16.4$, $p < .00005$, $MSE = 0.01$, but not for the remain

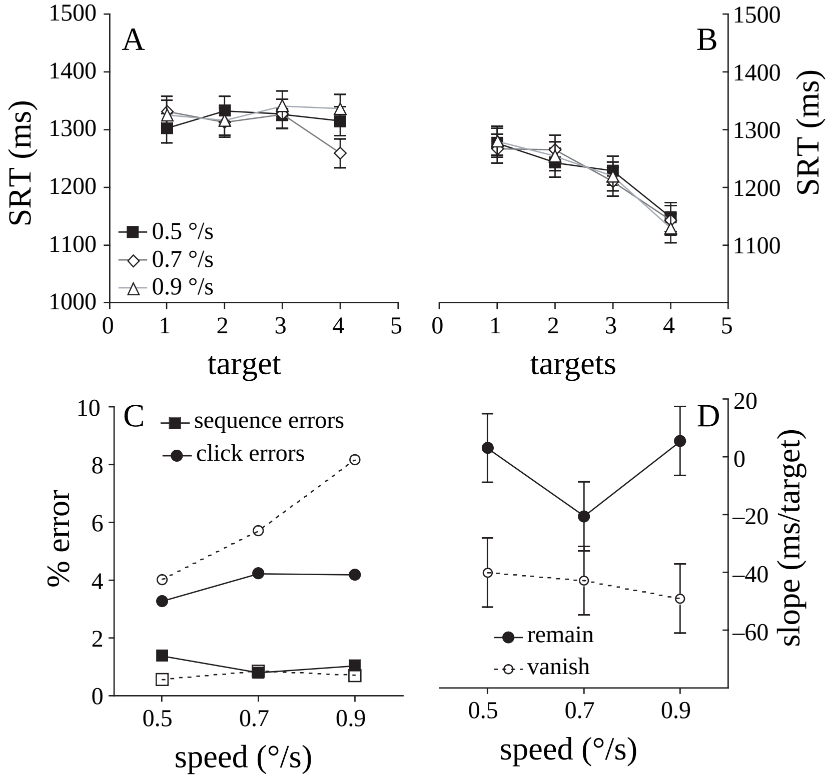


**Figure 6.** Data from Experiment 4. The top two panels plot SRTs as a function of speed and target number, with the remain condition observers shown in Panel A and the vanish condition observers shown in Panel B. Panel C depicts percentage error as a function of speed for the remain (filled symbols) and vanish (open symbols) groups. Note that error bars in this panel are hidden by the plotting symbols. Click errors are plotted as circles, sequence errors as squares. Panel D plots the slope of the SRT × Target functions for the remain (filled symbols) and vanish (open symbols) groups.

observers, $F(2, 20) = 1.6$, $p > .10$, $MSE = 0.01$. Sequence errors, however, did not respond to speed, $F(2, 42) < 1.0$, $MSE = 0.01$, nor was there a Speed × Group interaction, $F(2, 42) = 1.1$, $p > .10$, $MSE = 0.01$.

SRTs (Figure 6A and B) were analysed via a mixed factorial ANOVA with condition as a between subjects variable and speed and target number as repeated measures. The difference between the remain and the vanish groups was marginal, $F(1, 21) = 3.0$, $p = .10$, $MSE = 217{,}437$. There was a main effect of target number, $F(3, 63) = 19.7$, $p < .000001$, $MSE = 3990$, which interacted with condition, $F(3, 63) = 11.9$, $p < .000005$, $MSE = 3990$. There was no effect of speed, $F(2, 42) < 1$, $MSE = 6678$, nor did speed interact with target number, $F(6, 126) < 1$, $MSE = 6678$, or condition, $F(2, 42) < 1$,

$MSE = 3367$. The three-way interaction was not significant, $F(6, 126) = 1.5$, $p > .10$, $MSE = 3367$.

In order to further explore the Condition × Target number interaction, we computed the SRT × Target number slopes for the two groups (Figure 6D). Slope differed significantly between the two groups, $F(1, 21) = 23.9$, $p < .0001$, $MSE = 1151$, but did not vary as a function of speed, $F(2, 42) = 1.4$, $p < .10$, $MSE = 758$, nor did speed and group interact, $F(2, 42) = 2.0$, $p > .10$, $MSE = 758$.

## Discussion

The current data appear to replicate Experiment 3 at all three speeds. Notably, our prediction that slope should be steeper for slower speeds in the remain condition was not borne out. This would have produced a monotonic downward trend in the filled squares in Figure 6. Instead, we see a nonmonotonic function, with the shallowest slope at the slowest speed. This suggests that either setting the items in motion disrupts retrospective memory altogether, or the spatial representations of old targets has a very tight spatial gradient that has already fallen off substantially within the average response time in this experiment.

Consistent with the idea that click errors primarily reflect the difficulty of clicking on a moving target, click errors increased with speed in this experiment. This was particularly true for the vanish group. In this experiment, unlike in prior (and subsequent) experiments, we observed a difference between the remain and the vanish conditions in the click error rate. This might reflect a speed–accuracy tradeoff, such that observers in the remain condition were responding more slowly and cautiously than observers in the vanish condition.

Sequence errors were again quite low, unaffected by speed, and similar between the two groups, suggesting that the cognitive difficulty of the task was constant across all conditions.

# EXPERIMENT 5

In Experiment 5, we decided to further explore the spatial gradient idea by conducting one final speed manipulation in which object motion was present but minimized. That is, we moved the objects so slowly that there would be considerable overlap between original and final positions throughout the course of a trial. If retrospective memory is still disrupted under these conditions, it would seem to suggest that any change in position can render past information uninformative.

## Method

Twelve students from the University of Wales community participated in this experiment. The design was essentially the same as Experiment 3, with the following exceptions. First, speed was reduced to 0.25°/s. Screen resolution was set to $1152 \times 870$.

## Results

Error rates are shown in the right-hand panel of Figure 7. Neither click, $F(1, 8) < 1.0$, $MSE = 0.01$, nor sequence errors, $F(1, 8) < 1.0$, $MSE = 0.02$, differed as a function of condition.

The left-hand panel of Figure 7 plots SRTs as a function of condition and target. SRTs declined with target, $F(3, 30) = 4.2$, $p < .05$, $MSE = 5982$, and were faster in the vanish condition, $F(1, 10) = 8.5$, $p < .05$, $MSE = 7229$. These two factors interacted, $F(3, 30) = 6.2$, $p < .005$, $MSE = 3557$, such that the decrease in SRT with targets was observed primarily in the vanish data. Analysis of simple effects revealed that SRTs declined in the vanish condition, $F(3, 30) = 10.3$, $p < .0001$, $MSE = 4406$, but not in the remain condition, $F(3, 30) < 1$, $MSE = 5133$.

We can also analyse the interaction by looking at the SRT × Target number slopes, which were significantly steeper for the vanish condition ($-44.9 \pm 12.2$ ms/target) than for the remain condition ($-4.4 \pm 12.2$ ms/target), $F(1, 10) = 27.4$, $p < .0005$, $MSE = 329$.

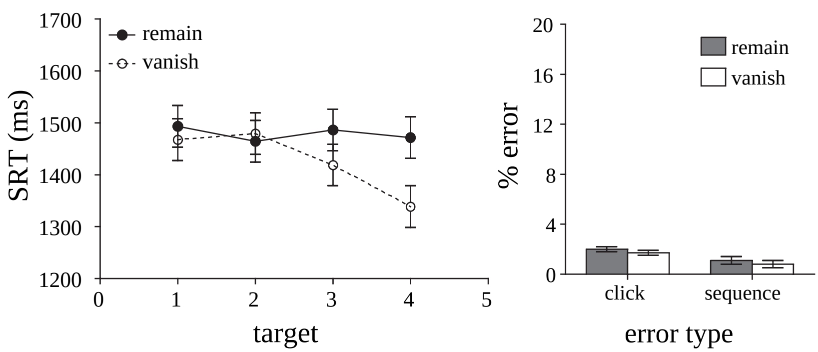


**Figure 7.** Left-hand panel plots SRT data from Experiment 5 as a function of target number. Filled symbols denote the remain condition, open symbols denote the vanish condition. Right-hand panel plots percentage error as a function of error type. Grey bars denote the remain condition, open bars the vanish condition.

## Discussion

Even at a very slow speed, when there was considerable overlap between original and final positions, retrospective effects were clearly absent. This suggests that any change of position, however small, can disrupt memory for old targets. Error rates in this experiment did not vary between conditions and were as low as those previously observed with static arrays (Thornton & Horowitz, 2004). We thus obtain the disruption of retrospective effects even in the absence of elevated error rates observed in previous experiments.

# EXPERIMENT 6

The evidence we have reported so far strongly argues against object-based coding of retrospective memory in the MILO task. However, our method for dissociating objects and locations is slightly different from that used in previous studies. Consider the work of Tipper and colleagues (Tipper, Driver, & Weaver, 1991; Tipper et al., 1994), which is perhaps most relevant here as it also used motion to dissociate location and objects. In a typical example of their paradigm, four objects might be presented, distributed symmetrically around fixation. One object is cued, then the objects rotate rigidly around the fixation point for 90°. At this point, one can probe either the originally cued object or the uncued object that is now in the originally cued location; the standard result is that both objects show some evidence of the effect in question, demonstrating that both object- and space-based reference frames are being used.

In the Tipper paradigm, the overall configuration of the display remains constant throughout the experiment, allowing the experimenter to dissociate *absolute* location from *relative* location. In contrast, in the studies described above, each item was given a unique and random motion path. This nonrigid motion pattern ensured that the overall display configuration was always disrupted.

In the current experiment, rather than having each item move on an independent trajectory, we yoked all items together and shifted them using a single random path. This rigid motion ensured that the display configuration was maintained throughout a trial, allowing observers to rely on relative locations. Does this manipulation allow observers to maintain retrospective memory for targets?

## Method

Sixteen observers from the Brigham & Women's Hospital Visual Attention Laboratory volunteer panel participated in this study. Experiment 6 was identical to Experiment 3, except that the motion of all items was slaved to a randomly chosen item and the overall display size was increased 18°. Speed was set to 1.4°/s. Display resolution was set to $1024 \times 768$.

## Results

Errors are shown in the right-hand panel of Figure 8. Neither click, $F(1, 15) < 1.0$, $MSE = 0.01$, nor sequence errors, $F(1, 15) < 1.0$, $MSE = 0.02$, differed as a function of condition.

## Discussion

The current data closely resemble those obtained in Experiments 3–5 (compare Figures 7 and 8). The rigid motion of the display seems to have speeded overall response times somewhat, and error rates fall between those seen with static arrays and random motion. However, there is still a very clear difference between the vanish and the remain conditions. This suggests that there is no retrospective advantage to being able to code past targets relative to each other or to an overall display framework.

# GENERAL DISCUSSION

We introduced the MILO task in order to study the spatiotemporal context of visual search. In our previous work, using static stimuli, we showed that both past actions and future plans affect responses to the current target

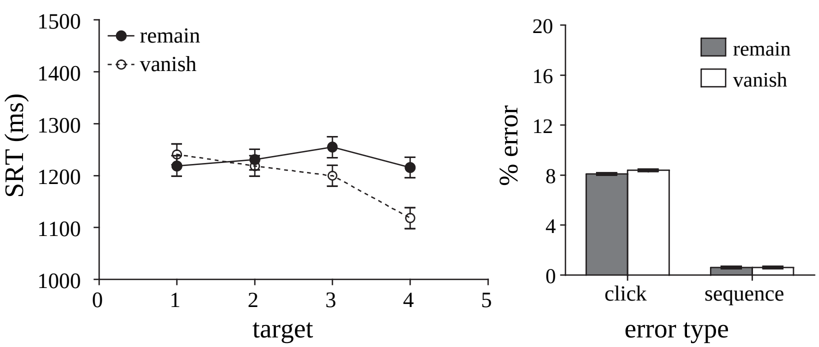


**Figure 8.** Left-hand panel plots SRT data from Experiment 6 as a function of condition and target number. Filled symbols denote the remain condition, open symbols denote the vanish condition. Right-hand panel shows percentage error as a function of error type and condition. Grey bars denote the remain condition, open bars the vanish condition.

(Thornton & Horowitz, 2004). In the experiments reported here, we used motion to explore the frame of reference within which these effects occur. Our results demonstrated a clear dissociation between the retrospective and prospective aspects of the task. Simply setting the search items in motion largely disrupted the utility of retrospective memory, while leaving prospective memory intact. These two aspects of the task will be discussed in turn.

The hallmark of retrospective effects in the MILO task is the near identical performance observed in the vanish and the remain conditions. With static arrays, a response to an item effectively removes it, improving search efficiency, regardless of that item's continued visibility. In all of the experiments reported here, motion eliminated this identity. That is, conditions in which items remained visible produced slower and generally less efficient search than when the item was physically removed. This was true even for very slow motion, where there was considerable overlap between initial and final locations (Experiments 4–5), and for relative motion, where the overall layout of the display remained constant (Experiment 6). We argue that this disruption demonstrates that the influence of past actions relies, to some extent, on a location-based frame of reference.

Clearly, this argument relies on the use of object motion to separate location-based and object-based reference frames. While this is a well-established strategy (Kahneman, Treisman, & Gibbs, 1992; Tipper et al., 1994), it is worth considering whether the addition of motion might also have affected performance in some other way. For example, suppose that motion simply increased the overall task difficulty. Could this have led to the observed divergence between vanish and remain conditions? In all of the current experiments, the overall level of response times were comparable to those obtained with static arrays (Thornton & Horowitz, 2004). Although we did note a slight rise in click errors (though not sequence errors), there was never a substantial difference *between* the vanish and the remain conditions, arguing against an underlying shift in task difficulty.

Could observers have accomplished performance levels comparable to those obtained with static arrays by strategically shifting limited resources away from memorizing prior target locations? Under this scenario, observers must have had sufficient remaining resources to support prospective memory in Experiments 1–2. In Experiment 3, however, we prevented observers from using prospective memory. If strategic shifting of resources were a feature of the current experiments, then observers might have been expected to free up resources devoted to prospective memory and shift them to retrospective memory. However, observers showed no more evidence for retrospective memory in Experiment 3 than in previous experiments. Finally, note that varying the speed of motion in Experiments 4 and 5 had no effect on retrospective memory. This lack of modulation is inconsistent with the notion that motion disrupts retrospective memory merely by increasing task difficulty.

Another more plausible alternative explanation for the impact of motion involves the overall density of the displays. Our measure of retrospective memory is the difference between the remain and the vanish conditions. Since the vanish condition is defined by a reduction in the number of items in the display over time, any advantage in comparison to the remain condition might conceivably be explained by reduced crowding effects (Pelli, Palomares, & Majaj, 2004). With static displays, this was not a concern, since the striking finding was that we observed no difference between conditions (Thornton & Horowitz, 2004). In the current experiments, we did observe an advantage for the vanish condition. Furthermore, while the interstimulus distances at the start of a trial were large enough to make crowding effects unlikely, the motion algorithm allowed objects to come closer, and even overlap.

Could crowding effects explain the difference between vanish and remain conditions in these experiments? Two findings suggest otherwise. First, slowing the speed of motion in Experiments 4 and 5, which would reduce the number of close encounters, had no effect on the vanish/remain difference. More tellingly, we still observed a difference in Experiment 6, in which the interstimulus distances remained constant. Crowding does not provide a very compelling explanation for the observed vanish/remain differences.

The most compelling explanation is that adding motion to the MILO task affects retrospective memory by disrupting location-based information. In the Introduction, we had predicted that both object-based and location-based coding might be used to support performance in the MILO task. Such dual coding has been reported for a number of other selection-related effects, such as negative priming (Shapiro & Loughlin, 1993) and IOR (McAuliffe et al., 2001; Tipper et al., 1994, 1999). Is there any indication that some form of retrospective advantage persists, thus hinting at some role for object-based coding?

In Experiments 4–6, the SRT function for the remain condition is essentially flat, with each subsequent response requiring an equal amount of time. Such a pattern indicates that the effective set size in the remain condition is constant, and that previously located targets are still interfering with ongoing responses. In these experiments, then, there appears to be no suggestion of residual retrospective coding.

In Experiments 1–3, however, even though there were clear differences between conditions, the remain slope was not flat. That is, responses occurring later in the sequence were speeded relative to earlier responses, even when items remained visible. Could this be evidence for residual object-based tagging? Although we cannot rule this out, there may be another explanation. We noted in our original paper (Thornton & Horowitz, 2004), that some component of the retrospective effect could be explained from a response competition point of view. Imagine that the motor system prepares a response program for each stimulus (Tipper, Howard, & Jackson, 1997; Welsh & Elliott, 2004). At the start of the trial, all of these programs compete to control the response, leading to long latencies. As responses are fired off, fewer programs are competing, and latencies decrease. Such a reduction in competition could thus give rise to non-zero remain slopes even in the absence of explicit item tagging.

Such an idea might also help explain the shapes of the SRT × Target functions. Visual search for characters in heterogeneous sets generally produces linear RT × Set size functions (Schneider & Shiffrin, 1977), so subtracting items should produce linear decreases in latency. However, in a number of the current experiments (1, 3, 5, and 6), as well as in our previous work (Thornton & Horowitz, 2004), even the vanish conditions produce nonlinear decreases. These nonlinear patterns may thus indicate the involvement of processes other than search itself. Selective reaching theorists are generally more concerned with trajectories than with latency, so their theories do not directly predict the SRT × Target functions. Nevertheless, response competition seems more likely to produce nonlinear functions than search.

What is still unclear in the current paper is why the remain slope is essentially flat in some experiments and not others. This may be evidence for some imperfect object-based inhibitory tagging or response competition. However, it is difficult to speculate further, since we can find no systematic difference between those conditions that produce some remain slope and those that produce none. The magnitude of the remain slope does not correlate with speed, as can be seen in Experiment 4. We also observed steep remain slopes both with (Experiments 3–4) and without (Experiments 1–2) the shuffle manipulation. Further research will be necessary to resolve this puzzle.

Motion clearly disrupts retrospective memory, and does so regardless of the speed of motion (Experiments 4–5) or the presence or absence of a global context (Experiment 6). Reducing the speed and introducing rigid motion improved performance (in both the RT and accuracy domains), but neither manipulation reinstated the ability to ignore old targets. Thus, the precise *location* of old targets seem to be quite crucial to performance. A straightforward interpretation of these data is that retrospective tagging of old targets occurs relatively early in visual processing, perhaps at the level of the salience map (Itti & Koch, 2001; Li, 2002) or activation map (Wolfe, Cave, & Franzel, 1989) thought to guide the deployment of attention in visual search. These maps are generally assumed to code locations, rather than objects, and retrospective memory could easily be implemented by inhibiting locations on such a map after a successful search (similar to how some models proposed to inhibit rejected distractor locations, e.g., Wolfe, 1994).

Interestingly, previous mechanisms that have been proposed to influence search by inhibiting old items on such maps, for example IOR (Klein, 1988; Klein & MacInnes, 1999; Shore & Klein, 2000) and visual marking (Watson & Humphreys, 1997), are also known to exploit object-based coding (McAuliffe et al., 2001; Tipper et al., 1994, 1999; Watson & Humphreys, 1998). Here, we found little evidence for such object-based coding. As the MILO task crucially involves both search and response components, this difference could reflect the additional involvement of the motor system. That is, the relevant maps may be more location-based because of the response component. Consistent with this idea, previous research has shown that response programs in cluttered displays are coded exclusively in location-based coordinates (Keulen, Adam, Fischer, Kuipers, & Jolles, 2002).

We turn now to the prospective aspects of the data. The most obvious indication of such planning in the MILO task is the large difference between the speed of response to an initial item compared to later items. Such a pattern is clearly visible in the data of Experiments 1–2 of the current paper and, as with static arrays, can be eliminated by introducing a shuffle manipulation (Experiment 3). Clearly then, motion does not eliminate prospective planning in the same way that it disrupts retrospective effects. This suggests that, in terms of future planning, items may be tagged and tracked within an object-based, rather than location-based frame of reference. For example, an observer may come across target $n+1$ during the search for target $n$. Target $n+1$ is recognized as a future target, but the observer cannot respond to it out of sequence. The object is therefore tagged as a future target so it can be easily recovered. As we suggested in the Introduction, such tagging might be accomplished via a spatiotemporal indexing mechanism such as Pylyshyn's (1989) FINSTs. Such indexes are intrinsically object based and would provide a natural substrate for this prospective memory.

In the current paper, we have focused more on the retrospective aspects of the MILO task. Clearly, an important direction for future research will be to more fully investigate the prospective aspects of the task. The sparing of prospective memory in dynamic arrays suggests a number of interesting research questions. With static arrays, we have shown planning effects in the MILO task that extend beyond the next item. We did this by leaving item $n+1$ or $n+2$, etc. stable, and shuffling further and further in advance. How far ahead might observers be able to track targets in the context of motion? Also, spatiotemporal indexing is known to be a capacity-limited mechanism (Trick & Pylyshyn, 1993). If we systematically vary the target set size, will we also see capacity limits on prospective memory?

In the typical MOT experiment, observers need to distinguish between two classes of moving objects: targets and nontargets. They usually do not need to discriminate among members of the target class. This task is thus more analogous to retrospective memory in our experiments, where the visual system needs to segregate old targets from other items. However, in prospective memory, observers presumably need to keep track of which upcoming target is which, since their responses must be ordered. Recent work suggests that capacity limits for this sort of task are more severe than on the segregation task.

Pylyshyn (2004) asked observers to track four targets out of eight moving objects. In one experiment, targets started off in different corners of the display, in another they had digits presented briefly at the start of the trial. In both cases observers could correctly discriminate targets from nontargets, without knowing which target was which. Horowitz et al. (in press) quantified this discrepancy. They asked observers to track four out of eight unique cartoon animals. At the end of the trial, the animals were masked by cartoon cactuses, and observers were either asked to indicate the locations of all the targets, or locate a particular target (e.g., the zebra). Observers could not adopt different strategies for different questions, since they did not know which question they would be asked until the end of the trial. Nevertheless, capacity was between three and four objects when observers were asked to segregate targets from nontargets, and one to two objects when asked to discriminate among targets.

Converging evidence comes from work with the multiple object permanence tracking paradigm (Saiki, 2003a, 2003b) and the object file reviewing paradigm (Henderson, 1994; Henderson & Anes, 1994). In multiple object permanence tracking (Saiki, 2002), a set of coloured shapes rotate behind a windmill-shaped occluder. The observer's task is to detect two objects exchanging colours. The change itself is obscured behind the occluder, so the observer must rely on object feature binding. Saiki finds that observers can maintain only one or two bound objects (2003a, 2003b). In object file reviewing experiments (Kahneman et al., 1992), observers see two displays: a preview display and a target display. Each display consists of a number of objects, such as boxes, containing letters. Typically, the preview and target displays are linked in such a way as to ensure that objects in the target display are seen as continuations of objects in the preview display. Observers are required to name the letters in the target display. Nonspecific priming occurs when letters from the preview show up anywhere in the target display, whereas object-specific priming occurs when a letter is presented in the same object in both displays. Henderson (1994; Henderson & Anes, 1994) showed evidence for both types of priming, and argued that they reflected two different representational systems, *object types* and *object files*, the latter being a limited capacity system capable of handling approximately two objects. Together, these results suggest that prospective memory could be quite limited.

Another important question for future research concerns response modality. We asked observers to move a mouse pointer. The results might have been different had we used saccadic responses. Fischer, Pratt, and Neggers (2003) have demonstrated that motor IOR is limited to the oculomotor domain, and not observed with pointing movements. This suggests that it may be useful to study the MILO task with real pointing responses and with oculomotor responses.

In any event, the apparent dissociation between prospective and retrospective memory suggested by the current experiments may provide a very powerful tool. That is, we can use motion to block retrospective effects, sparing prospective effects, and the shuffle manipulation to block prospective effects, sparing retrospective performance. Such a dissociation may allow us to shed further light on the nature of the mechanisms underlying these two aspects of the task using further behavioural or neuroimaging techniques.

## REFERENCES

Abrams, R. A., & Law, M. B. (2000). Object-based visual attention with endogenous orienting. *Perception and Psychophysics*, *62*(4), 818–833.

Alvarez, G. A., & Franconeri, S. L. (2005). How many objects can you track? Evidence for a flexible tracking resource [Abstract]. *Journal of Vision*, *5*(8), 641.

Cavanagh, P., & Alvarez, G. A. (2005). Tracking multiple targets with multifocal attention. *Trends in Cognitive Sciences*, *9*(7), 349–354.

Danziger, S., Kingstone, A., & Ward, R. (2001). Environmentally defined frames of reference: Their time course and sensitivity to spatial cues and attention. *Journal of Experimental Psychology: Human Perception and Performance*, *27*(2), 494–503.

Davis, G., Welch, V. L., Holmes, A., & Shepherd, A. (2001). Can attention select only a fixed number of objects at a time? *Perception*, *30*(10), 1227–1248.

Duncan, J. (1984). Selective attention and the organization of visual information. *Journal of Experimental Psychology: General*, *113*(4), 501–517.

Fischer, M. H., Pratt, J., & Neggers, S. F. (2003). Inhibition of return and manual pointing movements. *Perception and Psychophysics*, *65*(3), 379–387.

Ghafouri, M., Archambault, P. S., Adamovich, S. V., & Feldman, A. G. (2002). Pointing movements may be produced in different frames of reference depending on the task demand. *Brain Research*, *929*(1), 117–128.

Henderson, J. M. (1994). Two representational systems in dynamic visual identification. *Journal of Experimental Psychology: General*, *123*(4), 410–426.

Henderson, J. M., & Anes, M. D. (1994). Roles of object-file review and type priming in visual identification within and across eye fixations. *Journal of Experimental Psychology: Human Perception and Performance*, *20*(4), 826–839.

Horowitz, T. S., Klieger, S. B., Fencsik, D. E., Yang, K. K., Alvarez, G. A., & Wolfe, J. M. (in press). Tracking unique objects. *Perception and Psychophysics.*

Itti, L., & Koch, C. (2001). Computational modelling of visual attention. *Nature Reviews Neuroscience*, *2*(3), 194–203.

Kahneman, D., Treisman, A., & Gibbs, B. J. (1992). The reviewing of object files: Object-specific integration of information. *Cognitive Psychology*, *24*(2), 175–219.

Keulen, R. F., Adam, J. J., Fischer, M. H., Kuipers, H., & Jolles, J. (2002). Selective reaching: Evidence for multiple frames of reference. *Journal of Experimental Psychology: Human Perception and Performance*, *28*(3), 515–526.

Klein, R. M. (1988). Inhibitory tagging system facilitates visual search. *Nature*, *334*, 430–431.

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

Lamy, D., & Tsal, Y. (2000). Object features, object locations, and object files: Which does selective attention activate and when? *Journal of Experimental Psychology: Human Perception and Performance*, *26*(4), 1387–1400.

Li, Z. (2002). A saliency map in primary visual cortex. *Trends in Cognitive Sciences*, *6*(1), 9–16.

Loftus, G. R. M., & Masson, M. E. J. (1994). Using confidence intervals in within-subject designs. *Psychological Bulletin and Review*, *1*(4), 476–490.

McAuliffe, J., Pratt, J., & O'Donnell, C. (2001). Examining location-based and object-based components of inhibition of return in static displays. *Perception and Psychophysics*, *63*(6), 1072–1082.

Pelli, D. G., Palomares, M., & Majaj, N. J. (2004). Crowding is unlike ordinary masking: Distinguishing feature integration from detection. *Journal of Vision*, *4*(12), 1136–1169.

Pelli, D. G., & Zhang, L. (1991). Accurate control of contrast on microcomputer displays. *Vision Research*, *31*(7–8), 1337–1350.

Pylyshyn, Z. W. (1989). The role of location indexes in spatial perception: A sketch of the FINST spatial-index model. *Cognition*, *32*(1), 65–97.

Pylyshyn, Z. W. (2004). Some puzzling findings in multiple object tracking (MOT): I. Tracking without keeping track of object identities. *Visual Cognition*, *11*(7), 801–822.

Pylyshyn, Z. W., & Storm, R. W. (1988). Tracking multiple independent targets: Evidence for a parallel tracking mechanism. *Spatial Vision*, *3*(3), 179–197.

Rensink, R. A. (1990). Toolbox-based routines for Macintosh timing and display. *Behavior Research Methods, Instruments, and Computers*, *22*(2), 105–117.

Saiki, J. (2002). Multiple-object permanence tracking: Limitation in maintenance and transformation of perceptual objects. *Progress in Brain Research*, *140*, 133–148.

Saiki, J. (2003a). Feature binding in object-file representations of multiple moving items. *Journal of Vision*, *3*(1), 6–21.

Saiki, J. (2003b). Spatiotemporal characteristics of dynamic feature binding in visual working memory. *Vision Research*, *43*(20), 2107–2123.

Schendel, K. L., Robertson, L. C., & Treisman, A. (2001). Objects and their locations in exogenous cuing. *Perception and Psychophysics*, *63*(4), 577–594.

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

Shapiro, K. L., & Loughlin, C. (1993). The locus of inhibition in the priming of static objects: Object token versus location. *Journal of Experimental Psychology: Human Perception and Performance*, *19*(2), 352–363.

Shore, D. I., & Klein, R. M. (2000). On the manifestations of memory in visual search. *Spatial Vision*, *14*, 59–75.

Steinman, S. B., & Nawrot, M. (1992). Real-time color-frame animation for visual psychophysics on the Macintosh computer. *Behavior Research Methods, Instruments, and Computers*, *24*(3), 439–452.

Thornton, I. M., & Horowitz, T. S. (2004). The Multi-Item Localization (MILO) task: Measuring the spatio-temporal context of vision for action. *Perception and Psychophysics*, *66*(1), 38–50.

Tipper, S. P., Driver, J., & Weaver, B. (1991). Object-centred inhibition of return of visual attention. *Quarterly Journal of Experimental Psychology*, *43A*(2), 289–298.

Tipper, S. P., Howard, L. A., & Jackson, S. R. (1997). Selective reaching to grasp: Evidence for distractor interference effects. *Visual Cognition*, *4*(1), 1–38.

Tipper, S. P., Jordan, H., & Weaver, B. (1999). Scene-based and object-centered inhibition of return: Evidence for dual orienting mechanisms. *Perception and Psychophysics*, *61*(1), 50–60.

Tipper, S. P., Weaver, B., Jerreat, L. M., & Burak, A. L. (1994). Object-based and environment-based inhibition of return of visual attention. *Journal of Experimental Psychology: Human Perception and Performance*, *20*(3), 478–499.

Trick, L. M., & Pylyshyn, Z. W. (1993). What enumeration studies can show us about spatial attention: Evidence for limited capacity preattentive processing. *Journal of Experimental Psychology: Human Perception and Performance*, *19*(2), 331–351.

Watson, D. G., & Humphreys, G. W. (1997). Visual marking: Prioritizing selection for new objects by top-down attentional inhibition of old objects. *Psychological Review*, *104*(1), 90–122.

Watson, D. G., & Humphreys, G. W. (1998). Visual marking of moving objects: A role for top-down feature-based inhibition in selection. *Journal of Experimental Psychology: Human Perception and Performance*, *24*(3), 946–962.

Watson, S. E., & Kramer, A. F. (1999). Object-based visual selective attention and perceptual organization. *Perception and Psychophysics*, *61*(1), 31–49.

Welsh, T., & Elliott, D. (2004). Movement trajectories in the presence of a distracting stimulus: Evidence for a response activation model of selective reaching. *Quarterly Journal of Experimental Psychology*, *57A*(6), 1031–1057.

Wolfe, J. M. (1994). Guided Search 2.0: A revised model of visual search. *Psychonomic Bulletin and Review*, *1*(2), 202–238.

Wolfe, J. M., Cave, K. R., & Franzel, S. L. (1989). Guided Search: An alternative to the feature integration model for visual search. *Journal of Experimental Psychology: Human Perception and Performance*, *15*(3), 419–433.

*Manuscript received 2 May 2006*
*Manuscript accepted 25 October 2006*
*First published online March 2007*