Available online at www.sciencedirect.com
SCIENCE DIRECT®
Cognitive Psychology xxx (2005) xxx–xxx
Cognitive Psychology
www.elsevier.com/locate/cogpsych

Is motion extrapolation employed in multiple object tracking? Tracking as a low-level, non-predictive function [☆]

Brian P. Keane [*], Zenon W. Pylyshyn

Rutgers University Center for Cognitive Science, USA

Accepted 7 December 2005

UNCORRECTED PROOF

Abstract

In a series of five experiments, we investigated whether visual tracking mechanisms utilize prediction when recovering multiple reappearing objects. When all objects abruptly disappeared and reappeared mid-trajectory, it was found that (a) subjects tracked better when objects reappeared at their loci of disappearance than when they reappeared in their extrapolated trajectories, (b) disappearance episodes ranging from 150 to 900 ms had virtually no differential effect on performance, (c) tracking deteriorated monotonically as a function of displacement magnitude during disappearance, and (d) tracking did not depend on whether objects moved in predictable paths. Even objects that reappeared backward in their trajectories were tracked dramatically better than objects that appeared in their extrapolated trajectories. When all objects disappeared and reappeared in ways that implicated the presence of an occluder (i.e., with occlusion and disocclusion cues along fixed contours), tracking again was not predictive, and performance deteriorated with increased displacement. When objects reappeared predictably in 75% of trials, they were still tracked better when they reappeared at

---

[☆] Special thanks to Sharanya Chandrasekar and Amir Amirrezvani for help in programming, and Brian Scholl for suggesting the technique to make objects gradually occlude and disocclude. We also thank Todd Horowitz and David Fenscik for helpful discussions on object tracking. We additionally appreciate the input of the Visual Attention Lab, including Vidal Annan, John Dennis, Charles King, Alan Kugel, and Carly Leonard. Last, but not least, we are grateful for the advice of Rob Gordon and two anonymous reviewers. This paper constitutes the first author's Cognitive Science Certificate project and is supported by NIMH Grant 1R01-MH60924 to Z.W.P.

[*] Corresponding author. Present address: UCLA Department of Psychology, 1285 Franz Hall, Box 951563, Los Angeles, CA 90095-1563, USA.
E-mail address: keane@ucla.edu (B.P. Keane).

0010-0285/$ - see front matter © 2005 Published by Elsevier Inc.
doi:10.1016/j.cogpsych.2005.12.001

their points of disappearance. Theoretical implications of a non-predictive multiple object tracking mechanism are discussed.[1]

© 2005 Published by Elsevier Inc.

Keywords: Visual prediction; Motion extrapolation; Multiple object tracking; Displacement; Correspondence problem; Object persistence; Representational momentum; Flash-lag effect; Visual memory for location

1. Introduction

One of the most astonishing and important accomplishments of the visual system is that it allows us to see the world as having persisting objects. We might blink our eyes or make momentary saccades; objects might go behind and emerge from an occluder (Michotte, Thinès, & Crabbé, 1964/1991; Wertheimer, 1912), or appear and reappear to bring rise to apparent motion (Kolers, 1972). In all such cases, objects not continuously casting an image on the retinae are perceived to persist. How does this happen? This question is a version of what has been called the correspondence or temporal binding problem (Treisman, 1996; Ullman, 1985) and it brings with it two sets of questions (Scholl, Pylyshyn, & Franconeri, 1999). One set concerns what role, if any, that featural properties (e.g., shape, color, and size) play in allowing momentarily disappearing visual objects to be viewed as persisting. The other set concerns the role that spatiotemporal properties play in allowing objects to be viewed as persisting. The second set of questions will be of key interest in the present paper. In particular, we test the *prediction hypothesis*—that momentarily disappearing objects are most often perceived as persisting when they reappear with a velocity and a trajectory that was consistent with what was previously viewed. We will examine the prediction hypothesis from the viewpoint of tracking. If the prediction hypothesis is true, then objects will most likely be continuously tracked—and perceived to persist—when they reappear predictably relative to their pre-disappearance trajectories. Furthermore, the prediction hypothesis will be examined only when the spatiotemporal discontinuities are exogenously induced (i.e., when the distal stimuli disappear). Whether the hypothesis holds when subjects make the objects go out of view either by an eye-blink, saccade, or shift of the head or body will not be addressed.

In examining whether disappearing objects are tracked via prediction, we shall presuppose that attention is object-based. Though there is some evidence for space-based attention (e.g., Clark, 2000; Eriksen & Hoffman, 1973; Posner, Snyder, & Davidson, 1980), there is considerable psychophysical evidence (e.g., Baylis & Driver, 1993; Duncan, 1984; Egly, Driver, & Rafal, 1994; Kanwisher & Driver, 1992), neurophysiological evidence (O'Craven, Downing, & Kanwisher, 1999; Olson, 2001; Olson & Gettner, 1995; Roelfsema, Lamme, & Sprekreijse, 1998), clinical evidence (Behrmann & Tipper, 1994; Tipper & Behrmann, 1996), and conceptual arguments (Keane, 2005) that indicate that in at least some instances the fundamental units of attention are objects.

The experimental paradigm for addressing questions concerning the perception of object persistence and the prediction hypothesis is multiple object tracking (MOT). Originally introduced by Pylyshyn and Storm (1988), MOT involves picking out a target subset

---

[1] The following acronyms appear in this paper: ISI, interstimulus interval; MOT, multiple object tracking.

UNCORRECTED PROOF

(designated by a brief flash) of initially stationary visual objects, following the members of that subset for some duration as all objects on the screen independently move, and then re-identifying the targets with a mouse pointer at the end of a trial. It has been repeatedly shown that subjects can generally perform with at least 90% accuracy when tracking no more than five objects at a time (e.g., Cavanagh, 1999; Viswanathan & Mingolla, 1998; Yantis, 1992). MOT offers a useful paradigm with which to examine the prediction hypothesis because it allows a determination of how early visual tracking mechanisms operate in isolation from the influence of focused conscious attention. By understanding how momentarily absent objects can be referenced and tracked at the earliest levels, a better grasp can be had of how visual objects, and objects in general, can be viewed to persist at higher levels in perception and cognition.

The questions concerning predictive tracking will be addressed partly from the vantage of visual index (or FINST) theory (Pylyshyn, 1989). The early visual system, according to that theory, comes equipped with a series of pointers or indexes that can continuously reference a small number of visual objects in parallel. Indexes, or FINSTs (standing for Fingers of INSTantiation) as they are sometimes called, individuate objects to enable attentional processing (Pylyshyn, 1989; Sears & Pylyshyn, 2000), and they keep track of objects, despite changes in location or other physical properties (e.g., color, shape, and size). FINSTs are posited to explain, among other things, success in ordinary MOT. (For discussion on the array of visual capacities explained by visual indexes, see Pylyshyn, 2003.) Subjects are able to pick out the target items because the flashing of each target exogenously prompts the pointing of an associated index. Indexes follow their respective targets throughout their trajectories, according to FINST theory, in virtue of each index keeping track of its respective object's individuality. In this paper, we will clarify the role that visual indexes play in tracking across predictive and non-predictive spatiotemporal interruptions.

There are a number of reasons to suppose that object tracking may involve prediction. First, certain location illusions suggest it. In Freyd and Finke's representational momentum study (Freyd & Finke, 1984) observers tend to misidentify the offset location of a moving object as being slightly ahead of where it actually was. The tendency increased as a function of object velocity (Freyd & Finke, 1985) and predictability of traveled path (Finke & Shyi, 1988). In a version of the flash-lag illusion, a continuously visible segment of a rotating line appears ahead of a strobed segment of the same line (Nijhawan, 1994). Nijhawan concludes that an early visual mechanism adjusts for the spatial lag of a continuously visible (rather than flashing) moving object by extrapolating its instantaneous location (p. 257). Fu, Shen, and Dan (2001) showed that when two blurred, vertical bars move horizontally in opposite directions (left or right) and stop moving when they are perfectly aligned, subjects perceive the bars to have moved further in their trajectories than what they did.

Various other visual capacities appear to involve extrapolation. Palmer, Kellman, and Shipley (submitted), for instance, showed that when two object fragments appear at different points in space and time, subjects can extrapolate the contours of the first appearing fragment to be (roughly) in line with the contours of the later appearing fragment. Verghese and McKee (2002) showed that subjects automatically attend to locations immediately ahead of a moving target (see Fig. 1).

Though no studies, to our knowledge, have explicitly tested the prediction hypothesis for smooth linear motion of multiple objects, results from MOT occlusion experiments

UNCORRECTED PROOF

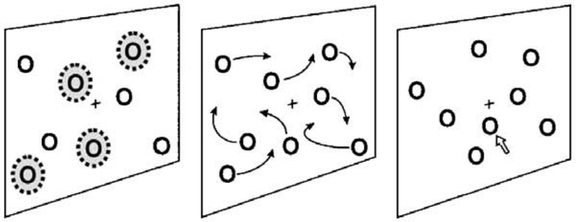

Fig. 1. In a standard multiple object tracking (MOT) trial, a subset of items will momentarily flash and begin to move for a short duration (usually between 5 and 15 s). At the end of a trial, subjects must identify the original flashed subset using a mouse. (Figure derived from Pylyshyn, 2000.)

(Scholl & Pylyshyn, 1999) suggest its validity (see Fig. 2). In the occlusion condition of that study, subjects tracked objects traveling behind screen-length, rectangular occluders. Even though objects momentarily moved completely out of view for an average of 225 ms, performance was high (never lower than 86 percent) and was no worse than when tracked objects (following the same trajectories) passed in front of occluders. The lack of performance decrement was found even when the occluders were invisible (or virtual), and even when different objects had different virtual occluders. This null result suggests that momentarily disappearing objects may be tracked predictively—that is, they may make use of pre-disappearance trajectory information to extrapolate to a probable locus of reappearance.

There also exist reasons to doubt the prediction hypothesis. Michotte and colleagues (1964/1991) showed that when a single object moves behind an occluder for a duration shorter than extrapolation would require, subjects will more likely perceive the same object exiting. Likewise, for a given exit–entrance time interval, when the displacement an object undergoes is less than it predictively ought to be, subjects are more likely to see the same object exiting the occluder (ibid). If the conditions under which we phenomenally perceive the persistence of momentarily disappearing objects are anyhow similar to

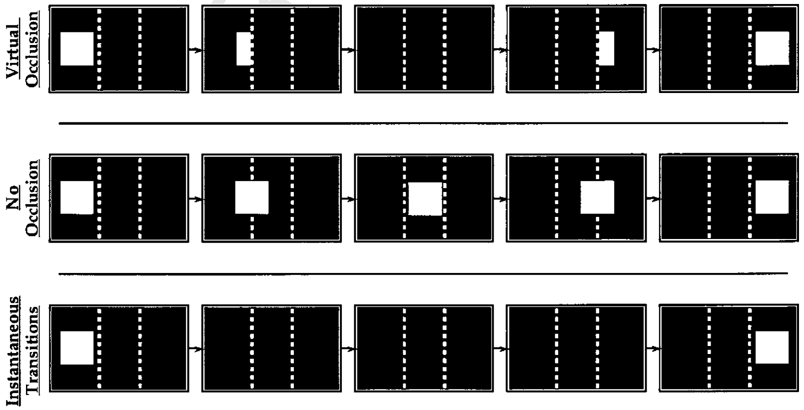

Fig. 2. Three conditions of Scholl and Pylyshyn's MOT study (1999, p. 268) that will be relevant to upcoming discussion. In the virtual occlusion condition, objects gradually disappear with cues of occlusion/disocclusion. In the no occlusion condition, objects do not disappear. In the instantaneous condition, objects disappear exactly when any portion overlaps with a virtual occluder.

those that are required for continuous tracking in MOT, then the prediction hypothesis looks unlikely.

Principles of apparent motion also provide evidence against the prediction hypothesis. According to Korte's "third law," increasing displacement will deteriorate the percept of persistence (when ISI is fixed), even when the increase leads to a predictive reappearance (Korte, 1915). If the correspondence process underlying apparent motion is similar to that underlying multiple object tracking, and if the visual system indeed prefers to minimize spatial or temporal separation at the cost of predictability, then the prediction hypothesis, again, cannot be correct.

Finally, Nijhawan's (1994) explanation for the flash-lag is by no means universally accepted. There are some who believe that the effect owes not to visual extrapolation of the continuously present stimulus but to a neural delay in the processing of the flashed stimulus (Whitney, Cavanagh, & Murakami, 2000; Whitney & Murakami, 1998). According to another view, the illusion owes to motion sampling errors (Brenner & Smeets, 2000). There exist various other theories of the phenomenon that do not require extrapolation (e.g., Eagleman & Sejnowski, 2000; Krekelberg & Lappe, 2000).

To investigate how the visual system keeps track of object identities across spatiotemporal interruptions and, in particular, to test the hypothesis that prediction is involved, we investigate a number of variations of the multiple object tracking task. All variations involve the disappearance of all objects about midway through a trial for some fixed duration, a blank screen interstimulus interval (ISI), and the reappearance of objects with some or no displacement from the loci of disappearance. In Experiment 1, we test, first, whether the kind of path (straight or curved) traveled before disappearance affects performance, second, whether objects are recovered better when they reappear in their extrapolated trajectories rather than where they disappear, and, third, whether increasing ISI degrades performance when objects reappear where they disappear. According to the prediction hypothesis, performance should be better for objects that travel in straight paths, better for predictive rather than zero-displacement reappearances, and ISI should have an effect when objects do not displace. In Experiment 2 we examine, first, whether objects are recovered better when they systematically displace along extrapolated paths by inappropriate amounts (either too much or too little), and, second, whether ISI affects performance when objects reappear displaced from their loci of disappearance. According to the prediction hypothesis, displacement along a trajectory (rather than no displacement) during a sufficiently long ISI should lead to better performance and ISI should have some effect on performance when objects displace. Experiment 3 examines whether objects are better recovered when they displace in their offset directions rather than in some other direction, and whether unpredictable changes in direction during disappearance disrupt tracking. According to the prediction hypothesis, objects that displace in their extrapolated trajectories and objects that retain their direction of movement should be recovered better than objects that do not. Experiment 4 is similar to Experiment 1, except that objects gradually (rather than instantaneously) move in and out of view. The prediction hypothesis predicts results similar to those of Experiment 1. Experiment 5 is also similar to Experiment 1, except that 75% (rather than 50%) of the trials involve predictive reappearances. This final experiment examines the extent to which practice can induce predictive tracking. The results accumulated over the course of the five experiments will ultimately give good reason to reject the prediction hypothesis, at least for the purposes of MOT. The primary factor in determining performance, instead, is the degree to which objects displace while out of view.

UNCORRECTED PROOF

2. Experiment 1: Examining effect of path, ISI, and displacement when objects reappear predictively or at loci of disappearance

Alvarez, Wolfe, Horowitz, and Arsenio (2001) showed that objects can be tracked at 70% accuracy when they all suddenly disappear and reappear in their trajectories. But it is unclear whether objects would be tracked with the same degree of accuracy if they reappeared in the "wrong" locations (i.e., at locations that do not correspond to where they would have been had they kept on moving while invisible). In Experiment 1, we consider this possibility. Objects will disappear for one of three intervals—150, 300, or 450 ms. For each of those intervals objects will either displace as if they had moved during the ISI or they will reappear where they disappeared. Furthermore, objects will either move in randomly curved paths for the duration of the trial or they will move in straight paths at constant speeds. If low-level tracking mechanisms are predictive, then subjects should perform better in MOT when objects reappear as if they had moved in their trajectories during the ISI. Moreover, if tracking is predictive, then performance should be better when objects travel in straight paths rather than in randomly curved paths. Representational momentum effects were found to increase with highly regular object paths (Finke & Shyi, 1988), so tracking might also improve with more regular object motion. Finally, if tracking is predictive, then the tendency to recover objects that reappear where they disappear should decrease with increasing ISI. Longer disappearance episodes should prompt a predictive tracking mechanism to search for objects farther away from the loci of disappearance.

2.1. Method

2.1.1. Subjects

Eleven Rutgers undergraduate students participated in a 50-min session for class credit. All subjects had normal or corrected-to-normal vision. No subjects were replaced and all finished the experiment in the scheduled time.

2.1.2. Apparatus

The tracking displays were presented on an iMac computer monitor with a resolution of 640 × 480 pixels and a refresh rate of 117 Hz. Subjects were positioned roughly 45 cm from the display monitor, creating a viewable screen that subtended an angle of 34° by 26°. All displays were programmed in VisionShell (Comtois, 2003).

2.1.3. Stimuli

Each trial involved four target objects and four non-target objects. With respect to their features, targets and non-targets were identical for all but the target-designation phase of a trial. Objects were white rings on a black background and their diameters subtended an overall angle of 2.7° of the viewable screen. The width of the annuli were 0.11°. Rings were used rather than solid circles since monocular T-junction depth cues minimize object-overlap tracking errors (Viswanathan & Mingolla, 1998, 2002) and since the aim is to isolate the effect of disappearance gaps on tracking performance.

Object trajectories were generated in real time during each trial, producing smooth and continuous motion. Each trial consisted of 590 static 8.55 ms frames producing 5 s of tracking. At the beginning of a trial, all objects were first shown as stationary on the

UNCORRECTED PROOF

display, flashed off and on five times and then began to move. The appearance of the non-targets did not change during the target designation phase of a trial.

To produce independent movement, objects were assigned random initial locations, directions, and (non-zero) speeds. Individual objects moved at 0° or ±0.053° or 0.11° (corresponding to 0 or ±1 or 2 pixels) in the x direction and y direction every other frame. Objects in the "curved" trials (described below) had "inertia" in that they retained their x or y velocity vector until an algorithm added or subtracted 0.053 to either vector with a probability set at .10. The object speeds in curved trials varied between 0 and 9.1°/s and the average object speed was 5.9°/s. Objects in the "straight" trials (described below) traveled constantly at an initially randomly selected value (±0.053° or ±0.11° every other frame) along the x direction and at one of the same speeds in the y direction. The straight-path object speeds averaged out to be 6.5°/s, and varied between 3.1 and 9.1°/s. When the edge of an object intercepted the edge of the tracking screen, the x or y velocity vector reversed its value, so that objects appeared to "bounce" geometrically off of the edge of the viewable screen.

2.1.4. Procedure and design

Subjects were seated in a darkened room in front of a monitor and operated a two-button mouse to perform the task. At the beginning of an experiment, subjects were given a demonstration and explanation of the multiple object tracking task. They were directed to attend to the blinking objects at the beginning of a trial, to follow those blinked objects for the duration of the trial, and to pick out those same objects at the end of a trial with a mouse. When objects were selected, their interior color changed from black to gray. Subjects were informed that there would be a screen blink-off for each trial, but they were not informed of the various manipulations tested. To motivate subjects to perform their best, at the end of each trial subjects were informed of the cumulative percentage of targets successfully identified for the block. At the end of a block, subjects received a prompt encouraging them to take a few minutes for break.

In all trials, all objects on the screen blinked off exactly once midway in a trial for 150, 300, or 450 ms. In half the trials, objects did not move during disappearance (the "non-move" condition); in the remaining trials, objects continued to move in their path (the "move" condition). An equal number of move and non-move trials were randomly distributed in each block and there were 40 trials per block. Blocks were individuated by both ISI (150, 300, and 450 ms) and trajectory path-type (straight or curved). A block of each duration of one path type appeared in the first half of an experiment, and a block of each duration of the alternative path-type appeared in the second half. Blocks in the first half of an experiment were arranged in a Latin square so that each ISI appeared at a different point in the experiment for different subjects. With respect to ISI, the sequence of blocks for the second half of an experiment repeated the block sequence of the first half. To balance for practice, every other subject began with a three block sequence of straight blocks. Preceding each experiment, a subject had six trials of practice, creating an experiment of 246 trials.

2.2. Results and discussion

Error rates were submitted to a 2 (Reappearance Type) × 3 (ISI) × 2 (Path Type) repeated measures analysis of variance. The type of reappearance was significant ($F = 157$,

UNCORRECTED PROOF

df = 1,10; $p < .001$) but not in the way that the prediction hypothesis entails. As Fig. 3 shows, subjects performed reliably better when objects reappeared where they disappeared than when they displaced in their trajectories. This held true even for the 150 ms block, for which there was little, if any, phenomenal difference between the non-move and move conditions (*t* test [two-tailed]: $p < .05$). The prediction hypothesis was also disconfirmed by the irrelevance of path type ($F = 2.12$, df = 1,10). An extrapolating mechanism better recovers disappearing objects when the pre-disappearance trajectories offer more reliable data on where an object will reappear. Straight-path trials obviously offered better data from which to extrapolate, but tracking mechanisms did not operate any differently in the presence of that data Fig. 4.

The undifferentiated performance in the non-move conditions across ISIs (as shown by paired, two-tailed *t* tests) also weakens the prediction hypothesis. For larger ISIs, a predictive mechanism will less likely recover objects that reappear with no displacement, but, again, no such effect was found.

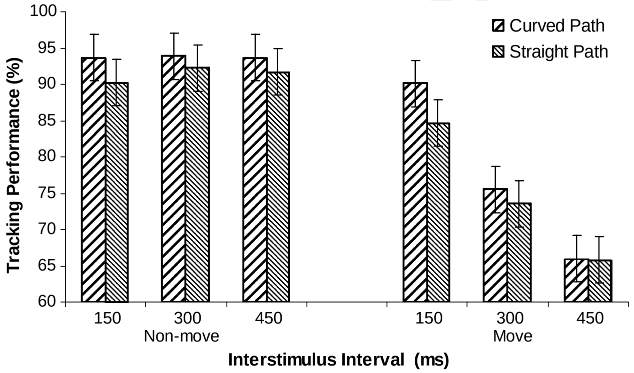

Fig. 3. Percentage of targets tracked and 95% confidence intervals for different ISIs, path-types, and reappearance types in Experiment 1. Confidence intervals in Fig. 1 and all other figures do not show between-subject variance and are calculated in accordance with the procedure of Loftus and Masson (1994).

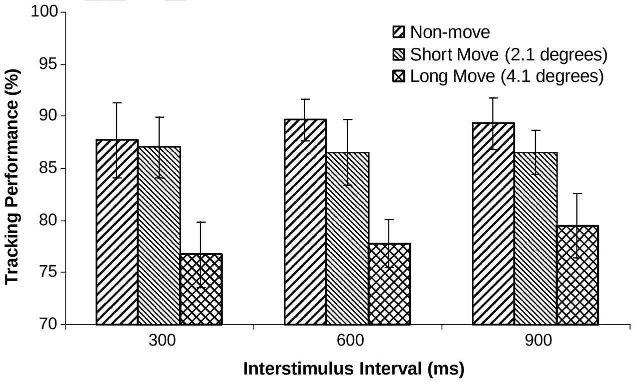

Fig. 4. Mean percent correct and 95% confidence intervals (with between-subject variance excised) for conditions in Experiment 2. Because the sphericity assumption is rejected, confidence intervals are calculated individually for each mean, as described by Loftus and Masson (1994, p. 484).

PROOF

While performance in the non-move conditions was not affected by ISI, performance in the move conditions was. Subjects tracked increasingly worse in the move condition relative to the non-move condition as ISI increased ($F = 72.9$, df = 2,20; $p < .001$). Even after the move and non-move data were pooled together, multiple pairwise comparisons (with a Bonferroni adjustment for multiple comparisons) showed that performance at each ISI significantly differed from every other ISI ($p < .004$). From the foregoing results, we conjecture that the drop off in performance in the move condition owes not to an increase in ISI per se, but to an increase in displacement that is associated with an increase in ISI. The view that performance depends primarily on displacement hereafter will be referred to as the *displacement hypothesis*; it will be relevant to the remaining experiments in the present paper.

3. Experiment 2: Effects of displacement and ISI on tracking

In the previous experiment, performance on the non-move condition was the same for all ISIs. It is possible that varying ISI for objects that displace a given amount also will not affect performance, in which case a stronger argument can be made that the visual system does not engage in prediction. Moreover, in the previous experiment, although subjects do worse when objects reappear in their extrapolated (predicted) location than when they reappear at their loci of disappearance, it is possible that the visual system will recover an object best when it displaces forward but reappears behind or ahead of its predicted location. For example, the visual system may prefer to recover objects that displace forward, but it may consistently mislocate those objects as being behind their predicted locations (Cooper & Munger, 1993). To test for these possibilities, we vary ISI and displacement independently so that objects reappear either ahead of, behind, or in their predicted locations at different ISIs.

3.1. Method

3.1.1. Subjects

Fourteen Rutgers University undergraduates participated in the experiment to receive class credit. All subjects had normal or corrected-to-normal vision. All subjects finished the experiment in the expected amount of time (about 50 min). One subject's data were eliminated because tracking performance was below the lower bound of the 99% confidence interval.

3.1.2. Apparatus and stimuli

The apparatus and the display in this experiment were the same as in the previous experiment except that in order to minimize the possible effect of phosphor decay, polarity was reversed so that objects were black rings on a white screen. Objects were of the same size and shape as Experiment 1.

3.1.3. Procedure and design

Objects traveled constantly along the x-axis either at ±0.11° or ±0.053° every other frame and constantly along the y-axis at one of the same values. In contrast to Experiment 1, the absolute value of the x-axis speed never matched that of the y-axis, so all objects traveled at the same speed—about 6.9°/s. There were three displacement conditions:

UNCORRECTED PROOF

"non-move" (which involved no displacement, as before), "short-move," in which objects displace forward in their path by 2.1°, and "long-move," in which objects displace forward in their path by 4.1°. Each of the three displacement conditions appeared in a block, block types were individuated by ISI, and there were three ISIs: 300, 600, and 900 ms. Within a block, half of the trials were non-move, and half of the trials involving displacement were short-move. Within a block, trials were randomized. Objects in the short-move condition appeared in their predicted location for the 300 ms ISI (since objects travelled 2.1° in 300 ms). Objects in the long-move appeared in their predicted locations for the 600 ms ISI (since objects travelled 4.1° in 600 ms). As before, there were 40 trials per block, and each block type appeared twice in an experiment. To balance for practice, blocks were arranged in a Latin square so that each block type appeared at a different point in the experiment for different observers. The combinations of the three blocks were then repeated, creating a total of six blocks in a trial. Preceding each experiment, a subject had six trials of practice, creating an experiment of 246 trials.

In previous experiments some objects in the move condition occasionally reflected off the edge of the screen, and therefore did not always move in straight paths, or maintain their pre-disappearance velocities for the duration of the ISI. To ensure that objects maintain their pre-disappearance line of motion through the ISI, we adopted the following design. We reduced the perimeter of the effective tracking area before offset in all conditions to accommodate for the greatest possible displacement along each dimension. Since in this experiment, objects could displace during an ISI by a maximum of 3.7° in either the *x* or *y* direction, the pre-disappearance tracking perimeter was reduced in length and width by 7.3° along each dimension. During regular tracking, objects stayed inside this reduced perimeter (bouncing off its edges if required). But if the trajectory during disappearance required the object to move outside the reduced perimeter, it would end up on the extrapolated straight line without having changed direction. Although such jumps outside the reduced perimeter were not frequent, this precaution ensured that the reappearance locations of objects were always consistent with its extrapolated (i.e., predicted) trajectory (see Fig. 5).

UNCORRECTED PROOF

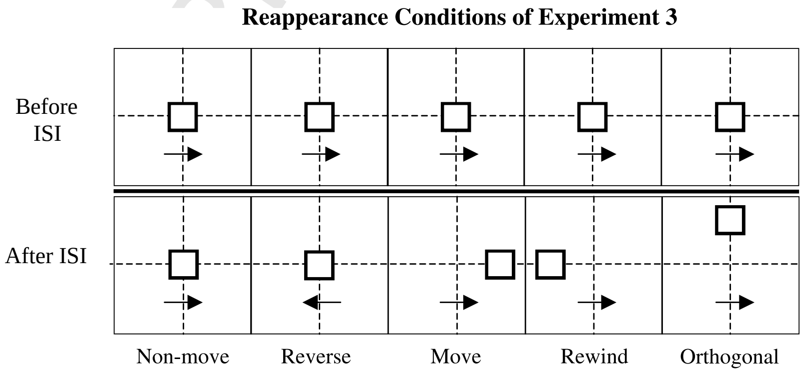

Fig. 5. Objects in the five conditions tested in Experiment 3. The cross-hairs represent the disappearance location, and the arrow represents the velocity vector of an object. (NB: An object in the orthogonal condition would also displace to the right of its trajectory on roughly 50% of trials.)

3.2. Results and discussion

We performed a 3 (Displacement) × 3 (ISI) repeated measures ANOVA. (Because Mauchly's test of sphericity was significant for both duration [$W = .482$, df = 2; $p < .02$] and displacement [$W = .270$, df = 2; $p < .001$], Pillai's trace is used in reporting those results.) As can be inferred from Fig. 4, tracking was completely insensitive to ISI ($F < 1$, df = 2,11). Paired t tests (two-tailed) between equivalent Displacement levels confirmed that ISI is irrelevant both when objects displace and when they do not. Consistent with Experiment 1, there was a main effect of Displacement ($F = 21.7$, df = 2,11; $p < .001$). Pairwise comparisons with a Bonferroni adjustment further showed that performance at each Displacement level differed significantly from performance at every other Displacement level, with larger displacements leading to worse performance ($p < .04$). These data confirm and strengthen the finding of Experiment 1, namely that performance depends primarily on the degree of displacement, and without regard to whether reappearances are predictive (see Fig. 6).

A surprising result was that subjects could still track at near 90% accuracy when objects disappeared for 900 ms. The memory involved in this task is too long to be sensory or iconic in character (Sperling, 1960), but it might be due to what Krekelberg (2001) calls "persistence of position." The capacity to recall position can help account for our findings and will be more closely considered in Section 7 (see Fig. 7).

4. Experiment 3: Is trajectory direction used to recover invisibly displaced objects?

In previous experiments it was shown that while subjects never track better than when objects reappear where they disappear, they can still track to some degree when objects reappear at their predicted location. For instance, in Experiment 1 performance was at 84% in the 150 ms move condition (with straight paths) and in Experiment 3 performance was at 84% in the 300 ms move condition. The limited success might owe, in part, to objects reappearing in their line of motion. The visual system might, so to speak, "prefer" to recover objects that do not displace over those that reappear in predicted locations, but it might also prefer objects that reappear in predicted locations over those that displace equally in other directions. We test this claim by comparing the move condition with an

RRECTED PROOF

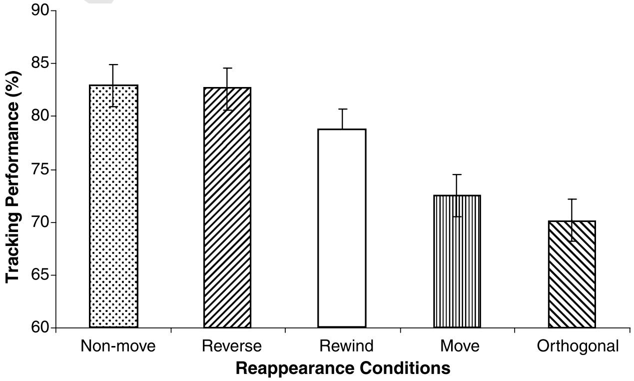

Fig. 6. Mean percent correct and 95% confidence intervals for reappearance conditions in Experiment 3.

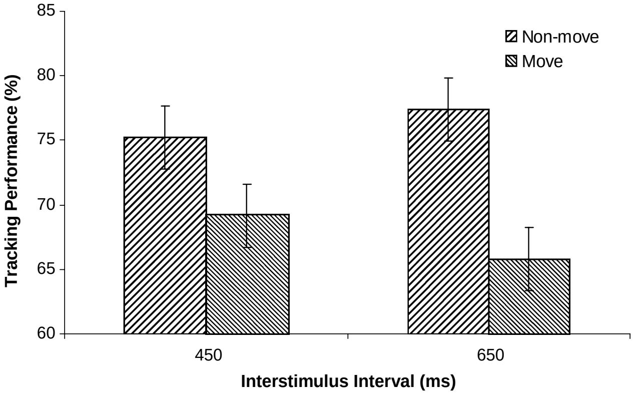

Fig. 7. Mean percent correct and 95% confidence intervals for disappearance conditions in Experiment 4.

“orthogonal” condition, in which objects displace perpendicular to their line of motion, and with a “rewind” condition, in which objects displace backwards to the line of motion. If tracking mechanisms are predictive, then performance should be worse in those conditions than in the move condition. On the other hand, if the displacement hypothesis is true, then performance in the move, orthogonal and rewind conditions should be undifferentiated (see Fig. 8).

A further way to examine whether tracking mechanisms utilize prediction is by introducing sharp discontinuities in direction of travel. In representational momentum studies, the degree to which subjects misidentify offset locations varies with an object’s expected trajectory, so that less predictable trajectories induce less forward shifting in identified locations (Finke & Shyi, 1988). To examine whether the direction of travel is encoded to recover reappearing objects in MOT, we introduce a new “reverse” condition in which object directions are reversed during the ISI. Equivalence of performance in the reverse and non-move conditions would suggest that continuity in traveled direction is not relevant for recovering reappearing objects. It would further suggest that displacement is the primary factor affecting correspondence.

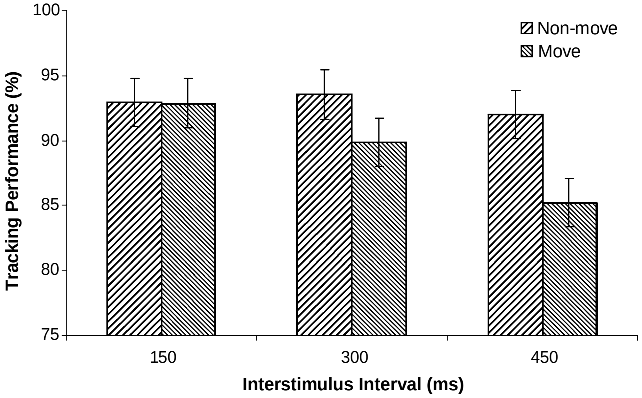

Fig. 8. Mean percent correct and 95% confidence intervals for conditions in Experiment 5.

4.1. Method

4.1.1. Subjects

Eleven Rutgers University undergraduates participated in the experiment to receive class credit. All subjects had normal or corrected-to-normal vision. No subjects were replaced and all finished the experiment in the scheduled amount of time (about 52 min).

4.1.2. Apparatus and stimuli

The testing apparatus was the same as Experiment 1. The display was the same as Experiment 2, except that objects were gray squares 1.1° on a side, and were outlined with a 0.11° black border. The squares retained their orientation throughout a trial with one edge remaining parallel to an edge of the screen. Because objects in all conditions travel at the same speed, objects in the move, orthogonal, and rewind conditions displaced the same amount (and not simply the same average amount).

4.1.3. Procedure and design

All conditions involved a 450 ms ISI. Two conditions were the same as in the previous experiment—move and non-move. Three new conditions were added. In the reverse condition, objects reappeared where they disappeared, but their directions were reversed. In the rewind condition, objects displaced in a direction opposite to their movement. In the orthogonal condition, objects displaced perpendicular to their path of motion. The direction of the perpendicular shift was randomized so that each object had an equal chance of shifting to the right or left of its line of motion. The move, orthogonal, and rewind conditions involved, displacements of 3.1° and velocity vectors did not change during that displacement. There were 50 trials per condition and trials from all conditions appeared in random order.

For all conditions, 2.05 s into a trial there was an invisible perimeter increase of 2.9° along each of the four sides of the screen, so that after the increase, the tracking area included the entire viewable screen. To ensure that objects in the move condition maintain a highly predictable linear motion through the ISI, we reduced the perimeter of the tracking area up to mid-trial offset to accommodate for the greatest possible displacement along each dimension. This was done in a manner similar to that used in Experiment 2, except that a small adjustment had to be made to make sure that the displacement in the move and rewind conditions was the same. In the rewind condition, objects disappeared exactly 2.5 s into a trial; in all other conditions, objects disappeared 2.05 s into a trial. Objects in the rewind condition disappeared after the screen perimeter increase so that they could displace in a straight path and with a magnitude equal to that of the move and orthogonal conditions.

4.2. Results and discussion

Bonferroni adjusted comparisons were drawn between each pair of conditions. Even though objects in this experiment were 60% smaller than Experiments 1 and 2, subjects still performed worse when objects reappeared in their extrapolated trajectories rather than with no displacement ($p < .001$). Furthermore, we found no difference between the non-move and reverse conditions. This finding obviously weakens the prediction hypothesis since a predictive mechanism should track objects better when the direction of travel does not radically change during the ISI.

UNCORRECTED PROOF

A third noteworthy result is the near equivalence between the orthogonal (mean = 70.2) and move (mean = 72.5) conditions. The lack of a difference is surprising considering that there was less variability in how objects reappeared in the move condition. Whereas in the move condition objects always moved forward relative to the line of motion, in the orthogonal condition objects could move right or left to their line of motion. It seems that tracking mechanisms are indifferent as to whether objects displace forward or to the side.

The prediction hypothesis, as we understand it, implies that objects reappearing further from their predicted locations will be tracked worse than objects reappearing closer to those locations. But this is not what we found. Performance in the rewind condition (mean = 78.8) was *better* than performance in the move condition ($p < .005$) and did not differ significantly from performance in the non-move (mean = 82.9) condition ($p > .2$). (This result was also reported at the Vision Sciences conference by Fencsik, Horowitz, Kliege, & Wolf, 2004; Keane & Pylyshyn, 2005). Once again, performance falls off as a function of the magnitude of displacement from the disappearance loci, though, surprisingly, performance falls off more slowly when the displacement is opposite to the line of travel. Possible reasons for the unexpectedly strong performance in the rewind condition are suggested in Section 7.

5. Experiment 4: Do occlusion cues encourage prediction?

Since objects in the world typically disappear through occlusion, tracking mechanisms may operate predictively if objects disappear in an ecologically valid manner (Gibson, 1979). In particular, if objects gradually go in and out of view as if they were moving behind and emerging from actual occluders (so that there are occlusion "cues"), then the visual system may utilize stored trajectory information to extrapolate to a probable reappearance location (Michotte et al., 1964/1991; Wertheimer, 1912).

Scholl and Pylyshyn (1999) put forth a related view in their tracking through occlusion study. In the "virtual occlusion" condition, square objects gradually slide behind and emerge from invisible occluders, which are twice the width of the objects (see Fig. 2). In their "instantaneous disappearance/reappearance" condition, an object blinks out of view exactly when any portion of it overlaps with the edge of the virtual occluder. Whereas subjects tracked at around 65% in the instantaneous condition, they tracked at an accuracy of almost 90% in the occlusion condition. The authors explain this finding

> (a) by appeal to the hypothesis that items can be tracked through interruptions in spatiotemporal continuity only when the interruption is perceived as being caused by the object moving behind an occluder, along with (b) the fact that instantaneous transitions (without accretion/deletion cues) do not implicate the presence of an occluding contour (p. 273).

The main claim must be changed in light of the presented results. Objects can be tracked with up to 93% accuracy when all objects instantaneously disappear and reappear at the same location. It may nevertheless be true that when objects displace during disappearance, tracking is enhanced when the occlusion is perceived as being caused by an occluder (Watamaniuk & McKee, 1995; Yantis, 1995). It may further be true that there is no such enhancement when objects do not displace at all while occluded.

To test whether occlusion cues allow for predictive tracking, all objects in Experiment 4 will begin to occlude at a specified time, go completely out of view, and begin to disocclude

UNCORRECTED PROOF

450 or 650 ms later. Objects either move for the duration that they are completely occluded (move condition) or they stop moving during that time (non-move condition). If subjects continue to track better in the latter condition, then cues of occlusion do not produce predictive tracking in MOT. Also, if subjects track equally well in the non-move condition for both ISIs, then that further disconfirms the prediction hypothesis, since larger ISIs should prompt a predictive mechanism to search for its respective target further away from the point of disappearance.

5.1. Method

5.1.1. Subjects

Eleven Rutgers University undergraduates participated in the experiment to receive class credit. All subjects had normal or corrected-to-normal vision. No subjects were replaced and all finished the experiment in the scheduled amount of time (about 42 min).

5.1.2. Apparatus and stimuli

The apparatus was the same as Experiment 1. The stimuli were the same as in the previous experiment, except that 2.05 s through a trial all squares gradually began to occlude. The gradual occlusion occurred such that for every pixel that an object moved along the *x*-axis, exactly one vertical line of object pixels disappeared (turned to white) from the leading edge of the object. In all trials, the occlusion duration was long enough for all objects to completely disappear. At a given point in a trial, for every pixel that a fully occluded object traveled along the *x*-axis, exactly one vertical line of pixels of the leading edge of the object became visible. In all trials, all objects completely disoccluded. The final effect was that each object appeared to fully move behind and emerge from its own virtual occluder, each of which was located and sized so that all objects began to occlude at the same time and to disocclude at the same time.

In order for occlusion durations of this experiment to be comparable to previous experiments, objects never entered occluders at excessively steep angles. Each object entered the occlusion at ±27° from the perpendicular or ±27° from the horizontal. This was ensured by: (a) assigning initial movement angles to be one of the eight directions that are exactly 27° from horizontal or vertical; (b) making objects travel in straight paths (as before); and (c) having objects bounce geometrically off of the tracking perimeter (as before). The fact that objects were squares of only 1.1° on a side also allowed us to minimize the amount of time objects spent occluding and disoccluding. With the foregoing initial conditions, all objects were fully occluded within 342 ms (40 frames) from the time that all objects began to occlude.

5.1.3. Procedure and design

Blocks in this experiment were individuated by ISI and there were two ISIs—450 and 650 ms. An equal number of move and non-move trials were randomly distributed in each block and there were 50 trials per block. A block of each duration appeared in the first half of an experiment, and that same sequence appeared in the second half of the experiment, creating a total of four blocks or 200 trials in an experiment (plus six practice trials at the beginning of the experiment). To balance for practice, every other subject began with the 450 ms block of trials.

UNCORRECTED PROOF

In the 450 ms block, all objects began to occlude at 2.05 s (t_1) into a trial, and began to disocclude at 2.5 s (t_2) into a trial. In the 650 ms block, all objects began to occlude at 2.05 s (t_1) into a trial and began to disocclude at 2.7 s (t_3) into a trial. In the move conditions, objects continued to move between t_1 and t_2 or t_1 and t_3 (depending on the block). In the non-move conditions, each object stopped moving exactly when it was fully invisible and started moving at t_2 or t_3. The contour of occlusion for all non-move trials, therefore, was exactly 1.1° along the x-axis from the contour of disocclusion. When objects resumed movement in the non-move conditions, they resumed the velocities had at t_1.

To ensure that all objects traveled in straight paths from t_1 to the time that they fully disoccluded, we used the same reduced pre-appearance screen technique adopted in Experiment 2. That is, in all trials there was a pre-disappearance screen reduction of 6.1° on each side of the screen, so that the total viewing screen before disappearance was 21.8° by 13.8°.

5.2. Results and discussion

Results from a 2 (Reappearance Type) × 2 (ISI) repeated measures ANOVA showed replication of previous findings. Objects were tracked better when they did not move during total occlusion versus when they re-appeared predictively ($F = 32.5$, df = 1,10; $p < .001$), performance differences between the move and non-move conditions increased for larger ISIs ($F = 8.26$, df = 1,10; $p = .017$), and ISI did not matter when objects did not displace (as revealed by paired, two-tailed t tests).

In contrast to Experiment 1, the drop off in performance in the move condition was insufficient to produce a main effect of ISI ($F < 1$, df = 1,10). The null result probably owes to the smaller displacement differences between the move conditions. Whereas the displacement in the longest ISI was roughly 200% the displacement magnitude in the smallest ISI in Experiment 1, the displacement in the 650 ms move condition (during total occlusion) was on average 100% larger than the displacement in the 450 ms condition. If displacement is measured from when objects go partially out of view to when they move completely in view, the displacement difference between the move conditions drops to 28%.

It seems, therefore, that even when multiple objects disappear and reappear gradually, there is no visual prediction. Tracking mechanisms, instead, keep track of disappearance locations and recover objects as a function of displacement from those locations.

6. Experiment 5: Controlling for frequency/expectancy effects of non-move trials

In Experiments 1 and 4, 50% of the trials involved reappearances at predicted locations, but in the other experiments the majority of the trials involved non-predictive reappearances. If objects usually reappear non-predictively, subjects might adopt the strategy of expecting them to reappear in non-predictive locations. Such a cognitive strategy might then account for the inferior performance on the move condition. To rule out this possibility, we run an experiment where 75% of the trials involve the move condition and the remaining trials involve the non-move condition. As discussed, if predictive tracking mechanisms exist, subjects should track better for the move condition than in the non-move condition at each ISI, and increasingly worse for the non-move condition as ISI increases.

UNCORRECTED PROOF

6.1. Method

6.1.1. Subjects

Twelve Rutgers University undergraduates participated in the experiment to receive class credit. All subjects had normal or corrected-to-normal vision. No subjects were replaced and all finished the experiment in about 50 min.

6.1.2. Apparatus and stimuli

The apparatus was the same as Experiment 1. The stimuli were the same as Experiment 1 except that all objects were black on a white background and moved in straight paths.

6.1.3. Procedure and design

The procedure and design were the same as Experiment 1 with the following differences. There were three different block types rather than six, corresponding to ISIs of 150, 300, or 450 ms. Each block contained 30 move trials and 10 non-move trials. To balance for practice, blocks were arranged in a Latin square so that each block type appeared at a different point in the experiment. The combinations of the three blocks were then repeated, creating a total of six blocks of trials for each subject.

6.2. Results and discussion

A 2 (Reappearance Type) × 3 (ISI) repeated measures ANOVA revealed results qualitatively identical to those found in Experiment 1. Subjects performed better when there was no movement versus when there was predictive movement ($F = 35.2$, df = 1,11; $p < .001$), performance differences between move and no-move increased with greater displacements ($F = 7.94$, df = 2,22; $p < .004$), and the drop off in performance in the move condition across ISI produced a main effect of ISI ($F = 8.23$, df = 2,22; $p < .003$). Furthermore, paired t tests (two-tailed) revealed no differences between the non-move conditions, indicating, once again, that the visual system does not take into account temporal interval when recovering multiple reappearing objects. Increasing the frequency and number of trials in which objects predictively move weakens, but does not eliminate, the pattern of results found before.

7. General discussion

The five experiments described here all suggest that under the conditions of MOT we explored, observers do not employ prediction when tracking multiple objects. When items disappear from view but keep moving, performance is impaired relative to a condition in which objects reappear where they disappeared. This result holds whether objects disappear suddenly or gradually with cues of occlusion and disocclusion. Contrary to the prediction hypothesis, we showed that for ISIs of at least 900 ms tracking performance drops off as a function of the magnitude of displacement without regard to the size of the ISI. Over a range of object sizes and movement speeds, the ability to track disappearing objects was found to be insensitive to predictability of traveled path, sudden changes in velocity, and displacement direction (provided that the direction is not backward). Even when 75% of trials involved predictive reappearances, and even when subjects could develop a strategy to visually predict, performance for the move condition still did not exceed that of the

UNCORRECTED PROOF

non-move condition. In what follows, we examine some issues raised by these findings and discuss how the traditional FINST account of MOT might need to be modified.

7.1. The unexpectedly strong performance in the rewind condition

Although increasing displacement leads to poorer tracking performance in all conditions, an unexpected finding was that much greater displacements are required to produce comparable impairment when objects reappear opposite to their direction of travel (the rewind condition). Specifically, in Experiment 3 we found no significant difference between the non-move and the rewind condition, despite a displacement of 3.1° in the latter. There are at least two possible reasons for this null result. One is that an implicit memory of the recently traveled trajectory improves observers' abilities to identify objects that fall on that part of the trajectory. That is, an object might be recovered more readily—via location priming (Maljkovic & Nakayama, 1996; Posner, Nissen, & Ogden, 1978)—when it reappears on its traveled trajectory. The view that there exists an implicit memory for trajectory is not new and indeed some suggestive evidence of it has been reported in MOT. Ogawa and Yagi (2002), for example, used the contextual cuing method (Chun & Jiang, 1998) to show that observers perform better in MOT if trajectories are repeated, even if the repetition is not noticed.

Superior performance in the rewind condition might alternatively owe to the misperception of onset or offset locations. The displacement hypothesis entails that smaller displacements improve tracking, so illusory mis-locations that reduce perceived displacements might also lead to better tracking. If the disappearance (offset) locations in the rewind condition were perceived as being behind where they actually were (relative to the direction of travel), the apparent displacement would be reduced. Likewise, if the reappearance (onset) locations were perceived as being ahead of where they actually were relative to previous movements, then the perceived displacement in the rewind condition would again be reduced. Unfortunately, offset localization errors will not help explain the rewind condition. When subjects tend to mis-identify offset locations, those locations are identified as being ahead of where they actually are, not behind (Freyd & Finke, 1984).

Forward onset mislocation errors have been reported and they might be a more likely cause of high performance in the rewind condition. In the Froehlich effect, for example, an object that abruptly appears is perceived as appearing further along its trajectory than what it really is (Froehlich, 1929). If the mechanisms responsible for the Froehlich effect were also causing subjects to misperceive where objects reappeared in MOT, then high performance on the rewind condition can be explained in terms of the smaller perceived displacement. One problem with this explanation is that the range of conditions in which the Froehlich effect appears is not well known. In the experiments we present, objects move between 3 and 9°/s, but the speeds typically used to demonstrate the Froehlich effect are more than 20°/s (Thornton, 2002). Also, location-shift phenomena have not been investigated with multiple targets moving in different directions at different velocities, as occurs in MOT. Tracking multiple objects differs from unitary attention-tracking in important ways, so we should not be surprised if findings of the latter do not apply to the former. Finally, and more seriously, there are some circumstances under which onsets can be perceived as being *behind* where they actually occurred, rather than ahead (Actis-Grosso & Stucchi, 2003). It would be premature, at best, to explain the rewind condition in terms of onset mislocalizations.

UNCORRECTED PROOF

Thus the reason for the superiority of the rewind condition remains unclear. It could owe to memory of a recently traveled trajectory, it could owe to a forward-shift of onset position, or it could owe to a more complex process that is not yet known. Our tentative conjecture is that some residue of the immediately past trajectory of a target remains when the target disappears and this results in priority being given to locations back from the point of disappearance over locations that are forward of the point of disappearance.

7.2. Visual memory and tracking momentarily disappearing objects

The principle that objects closest to the place of disappearance are favored when items reappear assumes that there is accurate storage of disappearance locations. It is doubtful that locations are stored iconically, since objects in our study disappeared much longer than the few hundred millisecond lifespan of iconic memory (Neisser, 1967; Sperling, 1960). A more likely kind of storage is a visual short-term memory or a VSTM (Phillips, 1974). A VSTM can store information about three or four objects concurrently (Cowan, 2001; Luck & Vogel, 1997) and can persist for durations on the order of seconds (e.g., Noles, Scholl, & Mitroff, 2005).

Although the capacity and duration limits of VSTM are within the range of the present studies, a storage more primitive than VSTM may be involved in the recall of object locations. On one hand, there is evidence that observers are poor at encoding multiple locations (indeed it has been argued that people can only accurately encode one location at a time, as measured by texture and alignment judgments [Hess, Barnes, Dumoulin, & Dakin, 2003]). On the other hand, there is evidence for a special form of position storage that encodes locations in parallel when visual objects disappear. Krekelberg (2001) has called this remarkable location memory "persistence of position" and has argued that it is independent of iconic memory and other forms of visual memory.

More direct evidence of the role of location memory in MOT comes from studies, such as those reported here, in which targets briefly disappear. As noted above, various researchers have shown that subjects could track multiple targets that were absent for 200 ms or longer (Alvarez et al., 2001; Scholl & Pylyshyn, 1999; Scholl et al., 1999). More recently, Horowitz, Birnkrant, Wolfe, Fencsik, and Tran (submitted) argued for a "task-switching" account of how observers manage to tolerate a gap in tracking, according to which observers "store the current task state whenever objects vanish" and then resume tracking based on the memory of where objects reappear. This account is very much in the spirit of the one we endorse and clearly involves accurate storage of target offset locations.

7.3. Can tracking mechanisms engage in prediction in MOT when disappearances are sequential?

An objection to our examination of the prediction hypothesis is that objects typically went out of view together at once. What would happen if objects disappeared sequentially one by one? Would visual prediction occur in that case? There are two reasons to think that it would not. First, in Experiment 4, although all objects began to disappear and reappear at the same time, they did not fully disappear or reappear at the same time. When some objects were completely out of view, other objects were half in view. Despite the disappearance/reappearance asynchrony, subjects continued to track objects best when they

UNCORRECTED PROOF

did not move during total occlusion. Second, results of other researchers (Horowitz, Birnkrant, Wolfe, Fencsik and Tran, submitted) show that for sequential disappearances in MOT, subjects perform better when objects reappear where they disappear versus when they reappear in their extrapolated trajectories. While other variations of the MOT task will have to be carried out in the sequential disappearance case (e.g., a reverse-sequential condition, an orthogonal-sequential condition, etc.), so far it looks like visual prediction does not occur when objects disappear one at a time.

7.4. Theoretical views on how momentarily disappearing objects are continuously tracked

As noted in the Introduction, visual index theory explains success in MOT by positing roughly four pointers that follow objects as they move about in the field of view. An index "sticks" to its respective object, on this view, not in virtue of the object retaining any particular feature, but in virtue of the object remaining continuously visible to the observer (e.g., Pylyshyn, 2001). Results from the present paper suggest that this traditional understanding of indexes must be changed. Despite disappearance intervals of up to 900 ms, visual indexes persist and seize the closest object that reappears within some nearby region. The further an object displaces from the locus of disappearance, the less likely that it will recover its associated index. Modifying FINST theory in this way is not unprecedented. In explaining infants' sensitivity to the numerosity of objects that move behind occluders, Leslie, Xu, Tremoulet, and Scholl (1998) also assumed that visual indexes can persist when objects become fully occluded. Future studies will have to determine the duration for which indexes can be sustained without a stimulus.

Notwithstanding the traditional view of indexes as purely causal mechanisms, there was speculation on whether indexes could retrieve and utilize motion information for extrapolation. In particular, it was postulated that a local-support predictor mechanism based on Kalman filters could predict object reappearances (Eagleson & Pylyshyn, 1989). Experiments discussed in the present paper give reason to doubt the Kalman filter model, at least for the conditions that we examined. Performance in MOT degrades rapidly with disappearance displacement, and neither the direction of displacement nor the continuity of motion mattered for performance (with the exception of the rewind condition discussed above). More importantly, in no case were observers able to track targets in the move condition better than in other conditions involving the same or less displacement. If early vision gathers trajectory information, that information cannot reliably be utilized to extrapolate in MOT.

Although the Kalman filter model and other prediction models (e.g., Marshall & Srikanth, 2000) appear to be inappropriate for characterizing how four objects are tracked, such models may still be useful for understanding the tracking of one or two items. They may be useful for understanding, for example, the tracking of individual dots that travel behind an occluder in a Brownian field (Watamaniuk & McKee, 1995), the interpolation of moving contours (Palmer et al., submitted), or the perception of misalignment of blurred bars (Fu et al., 2001). The task in future studies will be to identify more clearly the conditions under which visual extrapolation models will or will not be appropriate. Such inquiry, aside from being relevant to practical tasks like driving, will provide further insight into how we are able to see the world as having persisting objects.

UNCORRECTED PROOF

8. Uncited reference

Vinson and Reed (2002).

References

Actis-Grosso, R., & Stucchi, N. (2003). Shifting the start: Backward mislocation of the initial position of a motion. *Journal of Experimental Psychology: Human Perception and Performance, 29*, 675–691.

Alvarez, G. A., Wolfe, J. M., Horowitz, T. S., & Arsenio, H. C. (2001). Limits on multielement tracking [Abstract]. *Journal of Vision, 1*(3), 347a.

Baylis, G. C., & Driver, J. (1993). Visual attention and objects: Evidence for hierarchical coding of location. *Journal of Experimental Psychology: Human Percepton and Performance, 19*, 451–470.

Behrmann, M., & Tipper, S. (1994). Object-based visual attention: Evidence from unilateral neglect. In C. Umilta & M. Moscovitch (Eds.). *Attention and performance 15: Conscious and unconscious processing and cognitive functioning* (Vol. 19, pp. 351–375). Cambridge, MA: MIT Press.

Brenner, E., & Smeets, J. B. (2000). Motion extrapolation is not responsible for the flash-lag effect. *Vision Research, 40*, 1645–1648.

Cavanagh, P. (1999). Attention: Exporting vision to the mind. In C. Taddei-Ferretti & C. Musio (Eds.), *Neuronal basis and psychological aspects of consciousness* (pp. 129–143). Singapore: World Scientific.

Chun, M. M., & Jiang, Y. (1998). Contextual cueing: Implicit learning and memory of visual context guides spatial attention. *Cognitive Psychology, 36*, 28–71.

Clark, A. (2000). *A theory of sentience*. New York: Oxford.

Comtois, R. (2003). *Vision Shell PPC software libraries*. Cambridge, MA: Harvard Vision Laboratory.

Cooper, L. A., & Munger, M. P. (1993). Extrapolating and remembering positions along cognitive trajectories: Uses and limitations of analogies to physical motion. In N. Eilan, R. McCarthy, & B. Brewer (Eds.), *Spatial representations: Problems in philosophy and psychology* (pp. 112–131). Cambridge, MA: Blackwell.

Cowan, N. (2001). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioral and Brain Sciences, 24*, 87–114.

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

Eagleman, D. M., & Sejnowski, T. J. (2000). Motion integration and postdiction in visual awareness. *Science, 287*, 2036–2038.

Eagleson, R., & Pylyshyn, Z. W. (1989). *A computational model of a 2D ('FINST') tracking mechanism using spatio-temporal operators and a predictive filter*. London, Ont.: University of Western Ontario, Centre for Cognitive Science.

Egly, R., Driver, J., & Rafal, R. D. (1994). Shifting visual attention between objects and locations: Evidence from normal and parietal lesion subjects. *Journal of Experimental Psychology: General, 123*, 161–177.

Eriksen, C. W., & Hoffman, J. E. (1973). The extent of processing of noise elements during selective encoding from visual displays. *Perceptual Psychophysics, 14*, 155–160.

Fencsik, D. E., Horowitz, T. S., Kliege, S. B., & Wolf, J. M. (2004). Target reacquisition strategies in multiple object tracking [Abstract]. *Journal of Vision, 4*(8), 370a.

Finke, R. A., & Shyi, G. C.-W. (1988). Mental extrapolation and representational momentum for complex implied motions. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 14*, 112–120.

Freyd, J. J., & Finke, R. A. (1984). Representational momentum. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 10*, 126–132.

Freyd, J. J., & Finke, R. A. (1985). A velocity effect for representational momentum. *Bulletin of the Psychonomic Society, 23*, 443–446.

Froehlich, F. W. (1929). *Die Empfindungszeit*. Gustav Fischer.

Fu, Y., Shen, Y., & Dan, Y. (2001). Motion induced perceptual extrapolation of blurred visual targets. *The Journal of Neuroscience, 21*, RC172.

Gibson, J. J. (1979). *An ecological approach to visual perception*. Boston: Houghton-Mifflin.

Hess, R. F., Barnes, G., Dumoulin, S. O., & Dakin, S. C. (2003). How many positions can we perceptually encode, one or many? *Vision Research, 43*, 1575–1587.

Horowitz, T. S., Birnkrant, R. S., Wolfe, J. M., Fencsik, D. E., & Tran, L. How do we track invisible objects? Manuscript submitted for publication.

UNCORRECTED PROOF

Kanwisher, N., & Driver, J. (1992). Objects, attributes, and visual attention: Which, what, and where. *Current Directions in Psychological Science, 1*, 26–31.

Keane, B. P. (2005). *Visual objects: Philosophical and cognitive science perspectives*. Unpublished doctoral dissertation. New Brunswick: Rutgers University.

Keane, B. P., & Pylyshyn, Z. (2005). Does tracking disappearing objects in MOT involve predicting the locus of reappearance? [Abstract]. *Journal of Vision, 3*, 583a.

Kolers, P. A. (1972). *Aspects of motion perception*. New York: Pergamon Press.

Korte, A. (1915). Kinematoskopische Untersuchungen. *Zeizschrift für Pschologie, 72*, 193–296.

Krekelberg, B. (2001). The persistence of position. *Vision Research, 41*(4), 529–539.

Krekelberg, B., & Lappe, M. (2000). A model of the perceived relative positions of moving objects based upon a slow averaging process. *Vision Research, 40*, 201–215.

Leslie, A. M., Xu, F., Tremoulet, P., & Scholl, B. (1998). Indexing and the object concept: Developing 'what' and 'where' systems. *Trends in Cognitive Sciences, 2*, 10–18.

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

Luck, S., & Vogel, E. (1997). The capacity of visual working memory for features and conjunctions. *Nature, 390*, 279–281.

Maljkovic, V., & Nakayama, K. (1996). Priming of pop-out: II. Role of position. *Perception & Psychophysics, 58*, 977–991.

Marshall, J. A., & Srikanth, V. (2000). Curved trajectory prediction using a self-organizing neural network. *International Journal of Neural Systems, 10*, 59–70.

Michotte, A., Thinès, G., & Crabbé, G. (1964/1991). Amodal completion of perceptual structures (E. Miles and T. R. Miles, Trans.). In G. Thinès, A. Costall, & G. Butterworth (Eds.), *Michotte's experimental phenomenology of perception* (pp. 140–167). Hillsdale, NJ: Erlbaum.

Neisser, U. (1967). *Cognitive psychology*. New York: Appleton-Century-Croft.

Nijhawan, R. (1994). Motion extrapolation in catching. *Nature, 370*, 256–257.

Noles, N., Scholl, B. J., & Mitroff, S. R. (2005). The persistence of object file representations. *Perception & Psychophysics, 67*, 324–344.

O'Craven, K. M., Downing, P. E., & Kanwisher, N. (1999). FMRI evidence for objects as the units of attentional selection. *Nature, 401*, 584–587.

Ogawa, H., & Yagi, A. (2002). The effect of information of untracked objects on multiple object tracking. *Japanese Journal of Psychonomic Science, 21*, 49–50.

Olson, C. R. (2001). Object-based vision and attention in primates. *Current Opinion in Neurobiology, 11*, 171–179.

Olson, C. R., & Gettner, S. N. (1995). Object-centered direction selectivity in the macaque supplementary eye field. *Science, 269*, 985–988.

Palmer, E., Kellman, P. J., & Shipley, T. F. A theory of contour interpolation in dynamically occluded objects. Manuscript submitted for publication.

Phillips, W. A. (1974). On the distinction between sensory storage and short-term visual memory. *Perception & Psychophysics, 16*, 283–290.

Posner, M. I., Snyder, C. R., & Davidson, B. J. (1980). Attention and the detection of signals. *Journal of Experimental Psychology, 109*, 160–174.

Posner, M. I., Nissen, M. J., & Ogden, W. C. (1978). In H. L. Pick & B. J. Saltzman (Eds.), *Modes of perceiving and processing information* (pp. 137–158). Hillsdale, NJ: Erlbaum.

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

Pylyshyn, Z. W. (2001). Visual indexes, preconceptual objects, and situated vision. *Cognition, 80*, 127–158.

Pylyshyn, Z. W. (2003). *Seeing and visualizing: It's not what you think*. Cambridge, MA: MIT Press/Bradford Books.

Pylyshyn, Z. W., & Storm, R. W. (1988). Tracking multiple independent targets: Evidence for a parallel tracking mechanism. *Spatial Vision, 3*, 1–19.

Roelfsema, P. R., Lamme, V. A., & Sprekreijse, H. (1998). Object-based attention in the primary visual cortex. *Nature, 395*, 376–381.

Scholl, B. J., & Pylyshyn, Z. W. (1999). Tracking multiple items through occlusion: Clues to visual objecthood. *Cognitive Psychology, 38*, 259–290.

Scholl, B. J., Pylyshyn, Z. W., & Franconeri, S. L. (1999). When are featural and spatiotemporal properties encoded as a result of attentional allocation? [Abstract]. *Investigative Ophthalmology and Visual Science, 40*, 4195.

UNCORRECTED PROOF

Sears, C. R., & Pylyshyn, Z. W. (2000). Multiple object tracking and attentional processing. *Canadian Journal of Experimental Psychology, 54*, 1–14.

Sperling, G. (1960). The information available in brief visual presentations. *Psychological Monographs, 74*, 1–29.

Thornton, I. M. (2002). The onset repulsion effect. *Spatial Vision, 15*, 219–243.

Tipper, S. P., & Behrmann, M. (1996). Object-centered not scene-based visual neglect. *Journal of Experimental Psychology: Human Perception and Performance, 22*, 1261–1278.

Treisman, A. (1996). The binding problem. *Current Opinion in Neurobiology, 6*, 171–178.

Ullman, S. (1985). *The interpretation of visual motion*. Cambridge, MA: MIT Press.

Verghese, P., & McKee, S. P. (2002). Predicting future motion. *Journal of Vision, 2*, 413–423.

Vinson, N. G., & Reed, C. L. (2002). Sources of object-specific effects in representational momentum. *Visual Cognition, 9*, 41–65.

Viswanathan, L., & Mingolla, E. (1998). Attention in depth: Disparity and occlusion cues facilitate multi-element visual tracking. *Boston University Technical Report*, CAS/CNS-98-012.

Viswanathan, L., & Mingolla, E. (2002). Dynamics of attention in depth: Evidence from multi-element tracking. *Perception, 31*, 1415–1437.

Watamaniuk, S. N., & McKee, S. P. (1995). Seeing motion behind occluders. *Nature, 377*, 729–730.

Wertheimer, M. (1912). Experimentelle studien über das sehen von bewegung. [An experimental study of the visual perception of movement.]. *Zeitschrift für Psychologie, 61*, 161–265.

Whitney, D., Cavanagh, P., & Murakami, I. (2000). Temporal facilitation for moving stimuli is independent of changes in direction. *Vision Research, 40*, 3829–3839.

Whitney, D., & Murakami, I. (1998). Latency difference, not spatial extrapolation. *Nature Neuroscience, 1*, 656–657.

Yantis, S. (1992). Multielement visual tracking: Attention and perceptual organization. *Cognitive Psychology, 24*, 295–340.

Yantis, S. (1995). Perceived continuity of occluded visual objects. *Psychological Science, 6*, 182–186.

UNCORRECTED PROOF