Author's personal copy

Atten Percept Psychophys (2012) 74:691–702
DOI 10.3758/s13414-011-0265-9

A simple proximity heuristic allows tracking of multiple objects through occlusion

Steven L. Franconeri · Zenon W. Pylyshyn · Brian J. Scholl

Published online: 21 January 2012
© Psychonomic Society, Inc. 2012

Abstract Moving objects in the world present a challenge to the visual system, in that they often move in and out of view as they are occluded by other surfaces. Nevertheless, the ability to track multiple objects through periods of occlusion is surprisingly robust. Here, we identify a simple heuristic that underlies this ability: Pre- and postocclusion views of objects are linked together solely by their spatial proximity. Tracking through occlusion was always improved when the postocclusion instances reappeared closer to the preocclusion views. Strikingly, this was true even when objects' previous trajectories predicted different reappearance locations and when objects reappeared "too close," from invisible "slits" in empty space, rather than from more distant occluder contours. Tracking through occlusion appears to rely only on spatial proximity, and not on encoding heading information, likely reappearance locations, or the visible structure of occluders.

Keywords Attention · Attention: divided attention and inattention · Scene perception

Introduction

Disruptions are abundant in vision. Our eyes make ballistic movements multiple times per second, smearing the image of the world across the retinae. Movements of the eyes, head, and body dramatically change the visual input. And every blink shuts this input off entirely. Typically, we do not notice these disruptions, because our visual system blocks the motion signal produced by eye movements (e.g., Shioiri & Cavanagh, 1989), suppresses visual input during eyeblinks (e.g., Volkmann, Riggs, & Moore, 1980), and provides a representation of the environment that integrates multiple views (e.g., Henderson, 1997; Irwin, 1992). Other disruptions are due to the dynamic nature of the environment. Some types of objects tend to move, and it tends to be important to continuously monitor moving objects (such as cars, curious children, or soccer teammates). In the present experiments, we examined the ability to continuously attend to moving objects across a particular type of ubiquitous environmental disruption: occlusion, the temporary disappearance caused when an object passes behind another surface.

The ability to identify an object as the same individual across a period of occlusion can rely on several perceptual and cognitive processes, especially when only a single object is involved (see Scholl, 2007, for a review). But when multiple moving objects are present, the ability to track objects through occlusion presents special challenges and seems to recruit a special set of perceptual and attentional processes. The ability to continuously track moving objects is often studied using a multiple object tracking (MOT) task (Pylyshyn & Storm, 1988). Observers are shown a set of featurally identical objects on a computer screen, and a subset of these objects (e.g., four out of eight) are briefly marked as targets, by blinking or changing color. All of the (again identical) objects then begin to move unpredictably around the screen, and the observer's task is to mentally track the targets. After several seconds, the objects stop moving, and the observer identifies the targets. To succeed in this task, each object must be monitored continuously, and more than a moment's disruption can cause the observer to lose track of one or all of the target objects.

S. L. Franconeri (✉)
Northwestern University,
2029 Sheridan Rd,
Evanston, IL 60208, USA
e-mail: franconeri@northwestern.edu

Z. W. Pylyshyn
Rutgers University,
Newark, USA

B. J. Scholl
Yale University,
New Haven, USA

Author's personal copy

In the visual world, perhaps the most ubiquitous sort of environmental disruption is the occlusion of an object as it moves behind another surface. During MOT, however, brief occlusion of this sort does not impair performance; the visual system appears to treat such disruptions in terms of objects going in and out of sight, rather than as objects going in and out of existence—preserving the representations of those objects during the occlusion so that the tracking is uninterrupted (Scholl & Pylyshyn, 1999; cf. Gibson, Kaplan, Reynolds, & Wheeler, 1969). To track an object across occlusion, some properties of the original preocclusion behavior must be represented, in order to recognize the reappearing object as a tracked target. One straightforward mechanism for achieving this correspondence could be to store the last known location of the object. However, for disruptions that lead to larger position shifts (e.g., with wider occluders), the last known location of an object becomes increasingly unreliable as a guide to where it will reappear.

A role for extrapolation?

A solution to this problem would be to extrapolate the likely area of reappearance of an object given its last known trajectory. Several studies have tested this possibility, disrupting tracking displays either with occlusion or with abrupt disappearances and reappearances. Some of these studies suggest that extrapolation does not occur during MOT. For example, one study rendered all objects invisible for 100–500 ms during the tracking task while the objects continued on their extrapolated paths. Longer disappearance durations caused objects to move farther from their last seen positions, and this manipulation led to lower tracking accuracy (Horowitz, Birnkrant, Fencsik, Tran, & Wolfe, 2006). Control experiments showed that this drop in accuracy was due to the objects' shifted positions, and not to the temporal duration of the disruptions. A similar study also showed that the duration of invisibility did not affect performance: When objects were invisible for up to 900 ms but reappeared at their original location, accuracy was unimpaired (Keane & Pylyshyn, 2006). However, larger location shifts did impair tracking in this study: A shift of only 2.1° led to only a moderate tracking impairment, whereas a larger shift of 4° strongly impaired tracking (Keane & Pylyshyn, 2006). The fact that tracking performance was best when objects reappeared at their original disappearance locations suggests that no information about the previous trajectories of the objects was used to predict their future positions. In a more direct demonstration that extrapolation is not used during tracking, another experiment manipulated the moving objects such that each object either reappeared at its original disappearance position after a 450-ms disruption, 3.1° ahead (i.e., the predicted position given the disappearance time) or 3.1° in a direction perpendicular to its direction of motion (Keane & Pylyshyn, 2006). Performance was best for the original position, even though it was inconsistent with the predicted position. Performance was equally low for the extrapolated reappearance position and for the position perpendicular to the object's direction of motion. Thus, when target objects abruptly disappeared, there was no evidence for extrapolation beyond each object's last known location.

These studies might have failed to find evidence for motion extrapolation for two reasons. First, large numbers of targets disappeared synchronously. This is a rather different scenario than the types of disruptions discussed earlier: Whereas independent objects are frequently occluded by other surfaces in the natural environment at independent times (e.g., as each bird in a flock flies behind a tree, one by one), cases in which independent objects all disappear at the same time may be less common. Such situations may impose special challenges: Simultaneously predicting the future positions of each individual member of an entire frightened flock of independently moving birds might be hard to determine, while predicting the path of a bird or two at any given time might be more tractable. (In the initial demonstration of MOT through occlusion, each object was occluded at largely different times, as it happened to encounter static occluders while moving independently; Scholl & Pylyshyn, 1999).

In fact, there is some evidence that when the number of tracked targets is reduced, the motion direction of objects can begin to play a role in tracking. One study demonstrated such a dissociation by either showing observers a brief 2-s preview of object motion before the tracking task or showing them a static display of the last position from that preview. If performance were better in the motion preview condition, it would indicate that the motion information was helpful (Fenscik, Kleiger, & Horowitz, 2007). While there was no strong benefit when participants tracked a large number of targets, the motion preview helped performance when they tracked a small number of targets, suggesting that motion information could be used under low tracking load (see also Ellner, Flombaum, & Scholl, 2012). Another study showed that when participants were asked to track three targets (a relatively low load), there was some evidence for extrapolation when objects disappeared and they were required to click on the last known position of that object with the mouse (Iordanescu, Grabowecky, & Suzuki, 2009). The clicked position was more likely to be in a position within the 180° window of the object's "forward" motion path, rather than within the "backward" window.

A second reason that previous studies may have failed to find evidence for extrapolation is that, unlike occlusion, sudden disappearances and reappearances

Author's personal copy

may trigger a disruption to ongoing processing that is not well suited to extrapolation. In particular, the representation of object locations that is stored after a sudden display-wide disappearance may be an "offline" representation that cannot be updated easily, such that participants can perform difficult attentional tasks like visual search during the blank interval with little impairment to the tracking task (Alvarez et al., 2005). This representation may be available only when all objects disappear at once, such that object positions cannot be added incrementally over time or independently updated until the objects reappear (Horowitz et al., 2006; Keane & Pylyshyn, 2006).

The present experiments

We may summarize the work reviewed above by noting that there is only a small amount of evidence for extrapolation during MOT through disruptions but that most previous tests have employed some combination of ecologically unusual manipulations: sudden and synchronous disappearances and reappearances. In contrast, relatively little work has examined whether motion information is used to bridge disruptions in the case that is perhaps most common and where such extrapolation is most likely possible: under asynchronous gradual occlusion, with visible occluder borders that could predict an exact location of reappearance. Occlusion cues—specifically, the special types of gradual edge deletion and accretion associated with occlusion of moving objects—may provide visible "anchor" points for extrapolation. Such cues may signal to the visual system both *when* such extrapolation should occur (as the object starts gradually disappearing) and *where* the occluded object should reappear (at the far border of the occluder). Work in the contour interpolation literature supports the idea that powerful forms of extrapolation are possible for occluded objects, such that participants are sensitive to the alignment of visible edges with occluded (invisible) edges with surprising precision (Keane, Lu, & Kellman, 2007; Palmer, Kellman, & Shipley, 2006).

Previous research has signaled a special role during MOT for the gradual disappearance of objects that is characteristic of occlusion: When that cue is present, MOT through asynchronous disruptions is robust, but when it is disrupted, tracking is impaired. For example, when objects during MOT disappear via "implosion" (shrinking from all contours at once, rather than from a single contour adjacent to the occluder's border) and then later reappear via "explosion" (growing from a central point at the far occluder border), performance is greatly impaired relative to typical occlusion—even when the disappearances and reappearances occur at the same times, locations, and rates (Scholl & Pylyshyn, 1999). A further study showed that such impairments result from the implosion cues, rather than from the explosion cues: Performance was relatively impaired by implosion followed by normal disocclusion but was relatively spared by normal occlusion followed by explosion (Scholl & Feigenson, 2004). This suggests that gradual occlusion may be a special sort of cue for the visual system, and, indeed, this contrast between occlusion and implosion appears to also drive object persistence in several other contexts, including the tunnel effect (Flombaum & Scholl, 2006) and object tracking in infancy (Cheries, Mitroff, Wynn, & Scholl, 2008). In the present context, it is possible that this gradual occlusion cue not only supports continued tracking during MOT, but also does so by triggering some form of trajectory extrapolation.

Some past work has tested whether extrapolation is used when objects disappear via naturalistic occlusion and has yielded no evidence that occlusion leads to extrapolation (Horowitz et al., 2006; Keane & Pylyshyn, 2006). Our experiments have several properties that could provide a more sensitive measure of extrapolation. First, we use asynchronous occlusion instead of synchronous occlusion (as opposed to Keane & Pylyshyn, 2006). Second, both of these past studies used invisible occluders, which could impair an ability to predict the reappearance location of an object, whereas in most of the experiments presented here, we used visible occluders with predictable reappearance locations. Finally, we examined a wider array of circumstances where extrapolation might take place.

In the present experiments, observers tracked multiple objects through periods of asynchronous gradual occlusion at the borders of visible occluders, but we manipulated where and how the objects reappeared. In particular, objects could reappear (1) at locations shifted along the far occluder border from the nearest possible reappearance locations (Experiment 1), (2) at shifted positions that were consistent or inconsistent with the objects' extrapolated trajectories (Experiment 2), (3) while moving in new directions (Experiment 3), or (4) from locations that were "too close," from invisible "slits" in empty space, rather than from occluders' far contours (Experiment 4). To anticipate our results, these experiments did not reveal any substantial roles during MOT through occlusion for direction information, likely reappearance locations, or the visible structure of occluders; instead, tracking accuracy could be explained solely by the proximity of the disocclusion locations to the initial occlusion locations.

Author's personal copy

General method

Stimuli

Figures 1, 2, 3 and 4 depict samples of the occlusion manipulations in Experiments 1, 2, 3, 4a, 4b. Movies of all conditions are available at the first author's Web site (http://viscog.psych.northwestern.edu/demos/). All displays were generated on Apple eMac computers. Displays were created and controlled with custom software written in C using the VisionShell libraries (Comtois, 2004). Participants were seated approximately 55 cm from the monitor, which subtended 35.3° in width by 26.5° in height. Stimuli sizes are reported in pixels (640 × 480 resolution; 18 pixels ≈ 1° of visual angle). On each trial, tracking displays consisted of eight blue circles (20 pixels in diameter) containing white outlines (3 pixels thick) on a black background. There were also two occluders in Experiments 1 and 3, consisting of opaque black bars with a 2-pixel white outline, each 40 pixels wide and as tall as the display, placed 120 pixels (from each occluder center) to the left and right of the display center. There was only one occluder in Experiments 2 and 4a, 4b, which was wider (100 pixels) and centered within the display. Two-occluder conditions also contained a dim gray fixation ring (10 pixels in diameter) at the center of the display, although no fixation instructions were given.[1]

Objects moved at 138 pixels (7.6°) per second in Experiments 1 and 3. Because the wider occluders in Experiments 2 and 4a, 4b made tracking more difficult, object speed was slowed by 20%. Objects began moving in random directions and reflected off the display walls. On every video frame (138 Hz), object direction could be changed by X degrees, where $-0.23 < X < 0.23$ (with initial values randomly assigned from within this range), and with each new frame X increased or decreased, $-0.12 < \Delta X < 0.12$ (with the change value randomly chosen within this range). This level of inertia ensured that while object directions could change, they did not change so suddenly that motion information would not be diagnostic in predicting disocclusion locations. Although objects moved independently of each other and could, therefore, occlude each other, occlusion was minimized by generating each participant's trials in advance and retaining only those with low object occlusion rates (the lowest 5% of all trials generated, where, on average, object centers were 290 pixels apart, $SD = 136$).

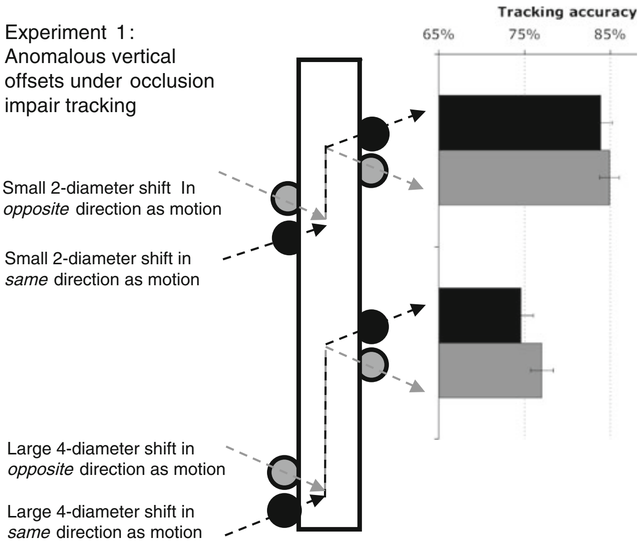


Fig. 1 In Experiment 1, objects entered the occluder and then shifted two or four item diameters up or down along the occluder, in either the same or the opposite direction as the vertical component of the item's trajectory. Results for the 100% manipulation conditions are depicted at right; results for the 50% conditions are discussed in the text. Error bars are within-subjects standard errors

Procedure and design

On each trial, the eight objects and occluder(s) appeared, and the target objects flashed (appeared and disappeared) 4 times, in 434-ms cycles. Objects then moved randomly (as noted above) and independently around the screen for 8 s. At the end of a trial, the participant was instructed to select the original four target objects with mouse clicks. As the mouse moved around the screen, the nearest unselected object was highlighted by changing the blue area to white, which became permanent when the participant selected the object.

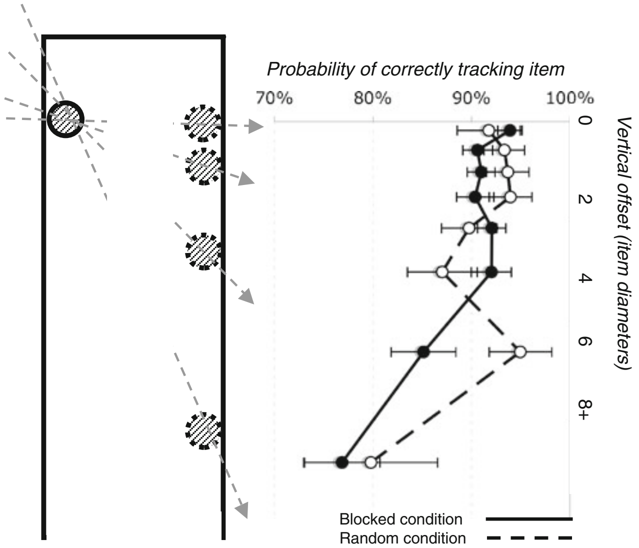
In most conditions across experiments, some objects underwent position changes during occlusion, as described in the Method section of each experiment. As soon as one of these objects became fully occluded, its vertical and horizontal position could change. In all conditions across each experiment, all objects exited the occluder immediately after entering the occluder, by "teleporting" to a location at the edge of its exit point, such that the object's outer border was barely visible by the subsequent video frame. While this manipulation meant that our displays did not mirror the way that natural occlusion events unfold over time, it was critical to ensure that manipulations of disocclusion locations were not confounded with the amount of time that the objects took to reach those locations. This "teleportation" manipulation goes surprisingly unnoticed, even by informed observers, and does not impair tracking performance (Scholl & Nevarez, 2002).

---

[1] Although the dynamics of eye movements during MOT are of interest (e.g., Fehd & Seiffert, 2008; Zelinsky & Neider, 2008), constraints on fixation seem not to affect performance. Pylyshyn and Storm (1988) eliminated trials on which subjects made eye movements, and they obtained results that were qualitatively identical to those of other studies that employed no special constraints or instructions concerning fixation (e.g., Intriligator & Cavanagh, 2001; Scholl et al., 2001; Scholl & Pylyshyn, 1999; Yantis, 1992).

Author's personal copy

a Experiment 2: Vertical shift under occluder impairs tracking, even when consistent with trajectory



b Experiment 2: Removing vertical shift aids tracking, even when *inconsistent* with trajectory

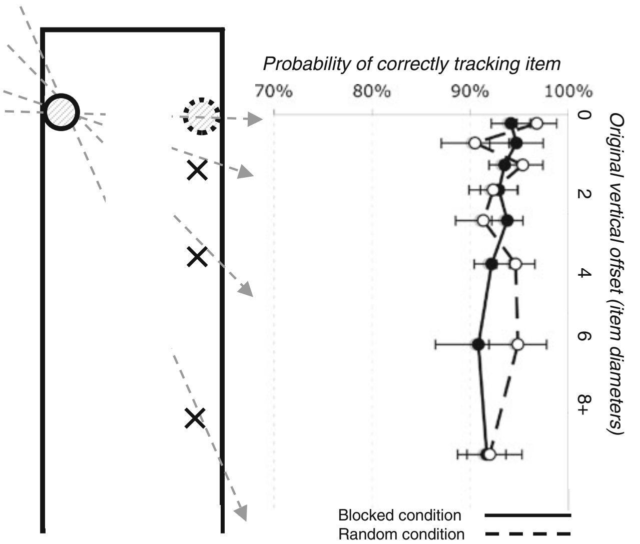


Fig. 2 **a** In the unaltered condition of Experiment 2, objects became occluded and reappeared at a location consistent with their motion history. The graph depicts tracking accuracy for a range of vertical shifts caused by the steepness of the entry angle. **b** In the same-vertical condition of Experiment 2, objects disoccluded at the same vertical position as they entered the occluder, regardless of their angle of entry. The graph depicts tracking accuracy for a range of would-be vertical shifts that would have resulted from the steepness of the entry angle. For both panels a and b, the dotted lines represent average accuracy for tracking individual objects in the random 50/50 condition that were either unaltered (panel a) or same-vertical (panel b). Error bars are within-subjects standard errors

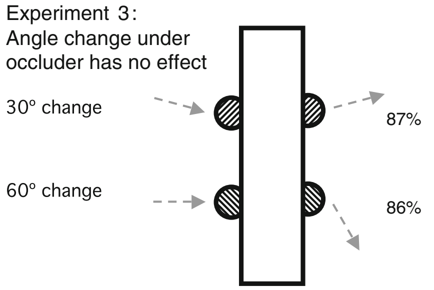
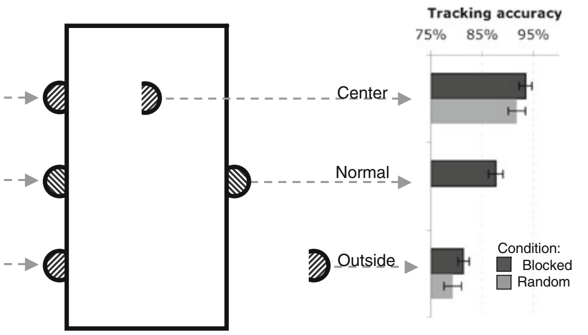


Fig. 3 In Experiment 3, tracking accuracy was identical for small 30° angle changes and large 60° angle changes that occurred while objects were occluded

Participants were given verbal instructions and two practice trials for each manipulation type. Occlusion manipulations were blocked, and participants were always fully informed of the manipulations used within each experiment. A dialog box appeared at the start of each new block containing a description of the manipulation (e.g., "In this block, the objects will shift vertically by a small amount (2 diameters) when they become occluded"). Data from the first trial of each block were discarded. In some blocks, all objects underwent

a Experiment 4: Horizontal shift away from entry point lowers accuracy



b Experiment 4: Controlling for object eccentricity

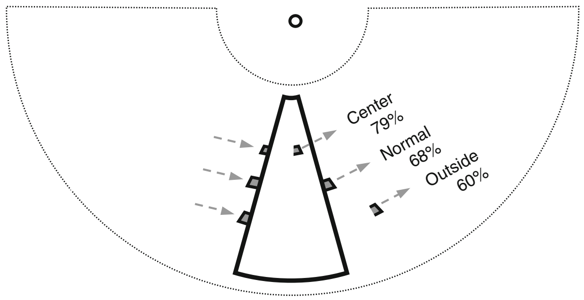


Fig. 4 **a** In Experiment 4a, objects exited at the occluder's center, at the opposite edge, or beyond the occluder's edge. The graph depicts average tracking accuracy across trials in the blocked conditions, as well as average accuracy for individual items in the random 50/50 condition. Error bars are within-subjects standard errors. **b** To eliminate a potential confound in Experiment 4a, Experiment 4b warped its tracking display around fixation. Despite eliminating the confound, the pattern of results was identical

Author's personal copy

the same manipulation upon occlusion. But to explore whether the observer's strategy could impact the target reacquisition process, each experiment also included a random (50/50) block consisting of a mixture of two manipulation conditions within each trial (e.g., "In this block, half of the objects will shift vertically by a small amount (2 diameters) when they become occluded, and the other half will exit normally"). Each manipulation included half of the targets and half of the distractors. Observers could not distinguish which manipulation an object would undergo until it became occluded. If the observers used a top-down strategy to reacquire disoccluded targets, this strategy should be harder to implement when the manipulations were not predictable. In contrast, if this manipulation had little effect, this would suggest that the underlying processing was more automatic and less strategic. For every experiment, the order of these conditions was counterbalanced across participants.

Experiment 1: location shifts along occluder borders impair tracking

Before testing whether tracking through occlusion is aided by prediction of an occluded object's future location, Experiment 1 first demonstrated that using asynchronous occlusion, objects that reappear farther from their original occlusion locations are harder to track. We asked participants to track four target circles among four distractors, while all items periodically (and asynchronously) became occluded behind one of two vertical bars. As an object moved behind an occluding bar, it underwent a vertical shift of either two or four item diameters (see Fig. 1). There were eight conditions, representing a full cross of the following three factors: (1) All objects underwent either two- or four-diameter vertical shifts while underneath the occluder; (2) this shift occurred for either 50% or 100% of the objects in a display on a given trial; and (3) the direction of the shift was in either the same or the opposite direction as the vertical component of the object's trajectory.

Method

Twenty-four undergraduates participated in return for course credit. When an object entered an occluder, its position was shifted vertically by two or four object diameters, depending on the current condition. The object’s angle of motion did not change after the shift. Motion paths were generated before the experiment, allowing rejection of trials where this shift would have moved the object outside the tracking area. There were four conditions: 100% of objects with two-diameter shifts, 100% with four-diameter shifts, 50% with two-diameter shifts (and 50% no-shifts), and 50% with four-diameter shifts (and 50% no-shifts). Each of these conditions was further divided into two types. In one type, the objects shifted vertically in the same direction as the vertical component of the object’s direction of motion (a shift “ahead”) or in the opposite direction (a shift “behind”). Blocks consisted of 10 trials of each of these eight condition types, for a total of 80 trials (plus practice trials). Before each new block, participants were told which of these eight conditions would make up the upcoming block.

Results and discussion

Two participants correctly selected fewer than three targets, on average, and were removed from the analysis. The results were submitted to a 2 × 2 × 2 ANOVA with factors of shift size (two or four item diameters), direction (in the same or opposite vertical direction as the object’s direction of motion), and rate (50% or 100% of occlusions were shifted). Accuracy data are depicted in Fig. 1, collapsed over shift rate. Accuracy was lower when shifts were four object diameters (M = 75.8%), relative to two object diameters (M = 84.5%), $F(1, 21) = 47.0$, $p < .001$, $\eta^2 = .69$. This finding confirms the results from past experiments using sudden disappearance instead of occlusion and shows that reappearance at larger distances from the original occlusion location leads to lower tracking accuracy. Accuracy was slightly lower when objects shifted in the same vertical direction as their motion (M = 79.3%), relative to the opposite direction (M = 81.0%), but this effect did not reach significance, $F(1, 21) = 2.9$, $p = .10$, $\eta^2 = .12$. This trend could reflect the fact that when a shift occurred in the same direction as the object's motion, the shift brought the object slightly farther from the original occlusion location by the time the object was fully disoccluded (see Fig. 1 for a depiction of this, and note how the black objects travel farther than the gray objects). Accuracy was also lower when shifts occurred for 100% of occlusions (M = 77.5%), relative to 50% of occlusions (M = 82.5%), $F(1, 21) = 25.7$, $p < .001$, $\eta^2 = .55$. There was also an interaction between the size of the shift and the proportion of trials on which shifts occurred (i.e., 100% vs. 50%), $F(1, 21) = 10.6$, $p = .004$, $\eta^2 = .34$, reflecting a larger accuracy difference between the two shift rates when the shift diameter was larger. There was no interaction between shift size and direction ($F < 1$).

In summary, when objects were occluded, introducing larger shifts from the original occlusion locations increasingly impaired tracking performance.

Author's personal copy

Experiment 2: location shifts impair tracking, even when consistent with an object's trajectory

In Experiment 1, a large shift along the disocclusion boundary (of four item diameters) away from the original location led to an accuracy drop of almost 10%, as compared with a small shift (of two diameters). These shifts were always anomalous, in that the disocclusion locations were never in the "correct" extrapolated positions. In this experiment, we tested whether similar impairments occur when an object's spatiotemporal history predicts that the object will disocclude far from the original occlusion location. We tested this in two ways. First, we compared performance in cases where objects always disoccluded at their extrapolated locations (Fig. 2a), but we analyzed performance on the basis of whether objects entered the vertical occluders while moving on near-horizontal motion paths (leading to small shifts from the objects' original occlusion locations), as compared with cases where objects entered at steeper angles (leading to larger predicted shifts if the trajectories were extrapolated). Second, we tested cases in which the objects always disoccluded at the closest possible point on the far occluder contour (Fig. 2b), and we analyzed whether performance in this situation depended on the angle of entry.

Method

Sixteen undergraduate students participated in return for course credit. When an object entered an occluder, its position either was left unaltered (i.e., was consistent with the object's continuing along a straight path through the occluder and exiting at the appropriate location; see Fig. 2a) or was manipulated such that the vertical component of the object's exit point was identical to that of the entry point (see Fig. 2b). There were three conditions: 100% unaltered, 100% same-vertical location, and 50% unaltered with 50% same-vertical. There were two blocks of each condition type, each containing 15 trials, for a total of 90 trials (plus practice trials). Block order was counterbalanced across participants.

Results and discussion

Two participants selected fewer than three targets, on average, and were removed from the analysis. Accuracy rates from the unaltered (*M* = 89.9%), same-vertical (*M* = 92.7%), and random 50/50 (*M* = 92.5%) conditions were submitted to a three-way ANOVA, which revealed a main effect of condition, $F(2, 26) = 5.0$, $p = .015$, $\eta^2 = .28$. This difference was driven by the unaltered condition having lower accuracy than the same-vertical condition, $t(13) = 2.5$, $p = .028$, and the random 50/50 condition, $t(13) = 3.5$, $p < .005$.

The lower performance for the unaltered condition suggests that performance is worse when objects disocclude at predicted locations, relative to "artificial" locations that are closer to the original occlusion locations. Because any trial in the random 50/50 condition contained two types of objects that underwent the two types of occlusion events, we then assessed performance according to individual objects, instead of individual trials. Overall accuracy for target objects in the random 50/50 condition that had an unaltered trajectory (*M* = 91.4%) was lower than that for objects that appeared at the same-vertical position (*M* = 94.0%), $t(13) = 2.2$, $p = .05$. In this random condition, accuracy was lower for unaltered trajectory objects than for altered trajectory objects (*M* = -2.6%), and this accuracy difference was similar to the difference between the two types of objects in the blocked conditions (*M* = -2.9%).

Unaltered object trajectories may lead to lower performance because each angled path leads the object to exit the occluder at a location that is farther from that object's entry point. If so, the performance decrement should be larger when the angle is steeper. To determine whether steeper occlusion angles led to lower performance, we recategorized the results of the unaltered condition within eight ranges of vertical change, with the constraint that each angle range have approximately the same number of observations. (This was approximately 200 observations for blocked conditions and 100 for each object type within the random 50/50 conditions. The bin for the largest angle for all conditions contained only approximately half those numbers.) Figure 2a depicts the average performance at each category of vertical change, as well as accuracy in the same-vertical condition. These data were submitted to a 2 × 8 ANOVA with condition (unaltered occlusions from the random 50/50 condition vs. blocked condition) and angle size as factors. There was no effect of condition or interaction between the factors (both *F*s < 1), but there was a decrease in accuracy as the magnitude of vertical change increased, $F(7, 91) = 5.3$, $p = .004$ with Greenhouse–Geissler correction, $\eta^2 = .29$.

Figure 2b shows the same analysis for the same-vertical condition, as well as for the same-vertical objects within the random 50/50 condition. For this analysis, the angle represents the object's original angle of entry and corresponding expected vertical offset upon disocclusion. For both the blocked and random 50/50 conditions, there was no effect of original angle on performance, all *F*s < 1.

As in Experiment 1, tracking performance was better when objects disoccluded closer to their original occlusion locations, as demonstrated by the higher performance in the same-vertical condition. Performance remained high in the same-vertical condition even when the objects' motion history predicted that they should disocclude at a location far from the artificially designated one. This advantage remained in the random 50/50 condition, where observers

Author's personal copy

did not know whether a particular object would disocclude at an unaltered or same-vertical location. These results are consistent with past studies showing that tracking accuracy is highest when objects reappear closer to where they disappeared.

It might be tempting to seek evidence for extrapolation by comparing these results with those of Experiment 1. While shifts of four item diameters caused a drop in accuracy in Experiment 1, accuracy did not significantly drop until objects had shifted eight diameters when the shift was predicted by the object's motion history in this experiment. This might appear to suggest that extrapolation attenuated the effects of shifts caused by moderate angles of entry. However, aside from the between-subjects comparison, the display conditions also differed in a crucial way between the two experiments: The occluders in Experiment 2 were 2.5 times wider, making the smallest possible shifts much larger, even with no vertical offset. Four diameters of vertical shift should have far less of an effect on performance when the "baseline" horizontal shift is already far larger. Thus, the results of both Experiments 1 and 2 suggest that location of disappearance is the primary factor linking pre- and postocclusion instances of objects during MOT through occlusion.

Experiment 3: angle changes during occlusion do not impair tracking

In order for extrapolation to occur, the angle of motion of an occluding object would need to be represented in order to perform the subsequent analysis of where it should disocclude. Experiment 3 explored whether an object's angle of entry can serve as a cue to that object's identity upon disocclusion (Fig. 3). If a target disoccludes while moving at a drastically different angle than when it became occluded, will tracking be impaired?

Method

Sixteen undergraduates participated in return for course credit. When an object reached the center of an occluder, its direction of motion either was held constant or was altered by 30° or 60°. There were four conditions: 100% 30° changes, 100% 60° changes, 50% 30° changes, and 50% 60° changes. There was one block of each condition, with 20 trials each, for a total of 80 trials (plus practice trials). Condition order was counterbalanced across participants.

Results and discussion

Data from 1 participant were lost due to computer error. The remaining accuracy data for each condition (100% 30° changes, M = 87%; 100% 60° changes, M = 86%; 50% 30° changes, M = 87%; and 50% 60° changes, M = 86%) were submitted to a 2 × 2 ANOVA with angle size change (30°, 60°) and rate (50%, 100%) as factors. There were no main effects and no interaction among these manipulations (all Fs < 1). Performance for conditions where objects changed by only 30° (M = 87%) was identical to performance for conditions where objects changed by 60° (M = 86%). These results suggest that the angle of motion itself is not used as a feature to link the preocclusion and postocclusion objects during MOT through occlusion.

Experiment 4a: proximity trumps visible occluder boundaries

Experiments 1 and 2 tested whether the motion direction of an object is used to predict its potential disocclusion location during MOT by varying vertical disocclusion position along the occluders' far borders. This experiment tested a complementary kind of spatial manipulation of disocclusion locations, more closely related to the structure of the occluders themselves. As was noted in the Introduction, one potential advantage of occlusion over sudden unexplained disappearances is that a visible occluder provides a salient cue as to an object's reappearance location: Even if you are not certain *where* along an occluder's far border an object will reappear, you can be fairly certain that it will appear *somewhere* along that border. In other words, the occluder's visible contours themselves may help to fuel a prediction about where an object will reappear.

In our displays, this possibility amounts to saying that the (vertical) occluders may at least afford a prediction about the horizontal component of an object's reappearance location: That location should be immediately adjacent to the occluder's far border. Is this true in practice? In other words, is MOT through occlusion best when the objects reappear at the closest possible points along the occluders far boundaries, or is performance best when the objects reappear at the closest possible points, regardless of the occluders' far boundaries? We tested this by varying whether objects gradually reappeared at, before, or beyond the far border of the occluder (Fig. 4a). If the structure of the occluder guides prediction, tracking accuracy should be highest when objects disocclude at the border. If not, tracking accuracy should be highest when the object disoccludes closer to the original occlusion location, even if that location represents an unlikely and unexplained exit point.

Author's personal copy

Method

Twelve undergraduates participated in return for course credit. When an object entered an occluder, its horizontal position was altered such that it exited the occluder from the occluder center (2.5 item diameters away, depicted in the top row of Fig. 4a), the normal exit location on the other side of the occluder (5 object diameters away, depicted in the middle row of Fig. 4a), or a location beyond the other side of the occluder (7.5 object diameters away, depicted in the bottom row of Fig. 4a). When the objects exited from a point at the occluder's center or outside the occluder, an invisible virtual occluder was placed over the disocclusion location so that the object would reenter the display with the same local accretion cues in all three cases (see Fig. 4). This invisible occluder existed only for disoccluding objects and did not affect the appearance of any other object. There were four conditions: 100% center disocclusion, 100% normal disocclusion location, 100% outside disocclusion location, and 50% center/50% outside. Blocks for the center, normal, and outside conditions each contained 9 trials, while blocks for the random 50/50 condition contained 13 trials. There were two blocks of each type, for a total of 80 trials (plus practice trials).

Results and discussion

Two participants correctly selected fewer than three targets, on average, and were removed from the analysis. Accuracy was higher for the center condition ($M = 93.5\%$, shown in the top row of Fig. 4a) than for the normal condition ($M = 87.8\%$, shown in the middle row of Fig. 4a), $t(9) = 2.5$, $p = .035$. Accuracy was similarly higher in the normal condition than in the outside condition ($M = 81.3\%$, shown in the bottom row of Fig. 4a), $t(9) = 2.9$, $p = .019$. Results in the random 50/50 condition were compiled by averaging accuracy across targets that had exited from the center ($M = 91.8\%$) and outside ($M = 79.3\%$). These accuracy rates were no different than in their corresponding blocked conditions, both $ts < 1.1$. In sum, tracking performance was unaffected by visible occluder boundaries. Instead, tracking performance was highest when objects exited as close as possible to the original occlusion location, even if that location was in the center of the occluder. Thus, the bridging of pre- and postocclusion instances of an object during MOT through occlusion seems not to depend on the visible contours of the occluders at all. This is consistent with the observation that MOT through occlusion is robust even if the occluders are themselves invisible—defined only via the gradual deletion and accretion cues that they impose (as in "virtual occlusion" manipulations)—even when each object has its own "private" virtual occluders in different locations (Scholl & Pylyshyn, 1999).

Experiment 4b: eccentricity control

Experiment 4a demonstrated that tracking performance is lower when objects disocclude farther from their original occlusion location, regardless of the visible occluder boundaries. However, in this experiment, the original and disocclusion locations were horizontally separated, with the fixation point in between these locations. Therefore, objects that were shifted farther from their original occlusion location also shifted farther from the fixation point. Because distance from fixation is known to amplify damaging crowding effects, including in MOT tasks (Franconeri, Alvarez, & Enns, 2007; Intriligator & Cavanagh 2001), it is possible that this amplified crowding effect was responsible for the lower performance with larger horizontal shifts. Experiment 4b removed this potential confound by warping the entire tracking display around the fixation point (Fig. 4b). In these new displays, greater "horizontal shifts" now led to greater angular changes, relative to fixation, but did not systematically alter the distance between an object and the fixation point.

Method

Nine undergraduates participated in return for course credit. Because all objects were squeezed into the lower visual field, the new displays were more crowded than in Experiment 4a. We therefore reduced the tracking load to three targets out of six total objects. All other aspects of the experiment were identical to those in Experiment 4a, except for the following display manipulations. Because the warping operation required a higher display resolution, this experiment used a resolution of 1,024 × 768 pixels at 75 Hz. Object sizes and speeds were adjusted accordingly to match the original displays before the warping operation was applied. To facilitate compatibility with our warping operation, prewarped objects were squares instead of circles. The display warping function took each point for each object (the four corners of each object and the occluder) and translated x-/y-coordinates into polar coordinates with x translating to angle, relative to fixation, and y translating to distance from fixation (with a 100-pixel buffer between the display's top and fixation and a total vertical extent of 360 pixels from top to bottom; 28.8 pixels $\approx 1°$ of visual angle). Postwarp, both objects and the occluder became segments of an annulus. (In Fig. 4b, this is more evident for the occluder, because the small size of the objects makes their curvature more difficult to distinguish.) There were three conditions: 100% center disocclusion location, 100% normal disocclusion location, and 100% outside disocclusion location. The mixed conditions from Experiment 4a were omitted (leaving 60 trials instead of 80).

Author's personal copy

Results and discussion

Results were similar to those observed in Experiment 4a. Performance was best in the center disocclusion location (*M* = 79%), lower in the normal disocclusion location (*M* = 68%), $t(8) = 3.6$, $p < .01$, and lower still in the outside disocclusion location (*M* = 60%), $t(8) = 2.5$, $p = .04$. After removing the potential confounding factor of distance from fixation, tracking performance was still strongly impaired by greater distance between the original occlusion and disocclusion locations, regardless of the visible occluder boundaries.

General discussion

We examined the information used to successfully track multiple objects through occlusion. Our primary conclusion is that this information is extremely sparse and, in particular, that it does not include details of the objects' motion histories. Instead, MOT through occlusion seems to rely on a simple heuristic based only on the proximity of reappearance locations to the objects' last known preocclusion locations (see also Keane & Pylyshyn, 2006). Experiment 1 demonstrated that targets that are displaced farther from their initial occlusion locations along the far borders of occluders are less likely to be successfully tracked. Experiment 2 showed that tracking is improved when objects reappear at those points along occluders' far borders that are as close as possible to the initial occlusion locations—even when those locations are grossly inconsistent with the extrapolated trajectories. Experiment 3 showed that although the motion trajectory of an object can be used to predict its future location, changing this motion angle suddenly upon disocclusion does not impair tracking. Finally, Experiments 4a and 4b showed that the visible structure of the occluding surface has no impact on how well an object is tracked: An object does not need to reappear at the opposite edge of an occluder in order to be recovered, and indeed, performance was best when objects reappeared closer to the original occlusion locations, even if that meant reappearing at "impossible" locations in the center of the occluder.

Overall, these results do not provide any evidence for the use of motion extrapolation in MOT through occlusion. This result is consistent with those of several previous studies (Horowitz et al., 2006; Keane & Pylyshyn, 2006). In the present study, this lack of extrapolation was observed during tracking through those disruptions that are perhaps most typical in the real world: gradual asynchronous occlusion and disocclusion with visible occluders. Instead, larger shifts of pre- and postocclusion locations always impaired tracking accuracy, regardless of whether such shifts were consistent with objects' trajectories.

These results are consistent with models of object tracking that focus on the representation of the positions of tracked objects within a segmented spatiotopic map of the world. Once a set of target object locations is represented as peaks of activation within this map (see Awh & Jonides, 2001), simple winner-take-all circuits (Koch & Ullman, 1985; Pylyshyn, 2000) could keep these peaks active while also shifting their position within the map in response to changes in an object's position within the segmented representation of the outside world (Franconeri, Bemis, & Alvarez, 2009). If these activation bumps moved only in response to the movement of the objects that caused them, they might remain in the last locations where the objects had been visible before a period of occlusion, rather than moving along with extrapolated paths. One study provided evidence for this idea by measuring detection accuracy for flashed probe dots on occluders during MOT through occlusion. When an occluder contained a tracked (but invisible) object, detection accuracy improved, suggesting attentional selection of that location (Flombaum, Scholl, & Pylyshyn, 2008), even with no visible object. Moreover, probe detection on objects was actually *better* when they were occluded, as compared with when they were visible (an "attentional high beams" effect). This could result from the need to store and maintain the last known positions of occluded objects, so that those locations can be used to help reacquire targets when they reappear. This type of location-based representation is consistent both with impairments caused by location shifts (as observed in Experiments 1 and 2) and with the lack of any impairments due to sudden postocclusion direction changes (tested in Experiment 3) or violations of visible occluder structure (tested in Experiments 4a and 4b). The most significant performance-limiting factor in tracking through occlusion thus appears to be the uncertainty associated with the representation of preocclusion locations.

We argue that location information is of primary importance for MOT performance. The most significant performance-limiting factor in *any* form of MOT may be the uncertainty associated with represented object locations, relative to the proximity of other objects that might steal the spotlight of activation and foil ongoing correspondence computations (Franconeri, Jonathan, & Scimeca, 2010; Franconeri, Scimeca, & Jonathan, 2012; Franconeri et al., 2008; Ma & Huang, 2009; Vul, Frank, Alvarez, & Tenenbaum, 2009). Our results cannot rule out a role for motion extrapolation across other types of disruptions or for tracking performance without disruptions. Some studies do show evidence of extrapolation in tracking tasks that use smaller numbers of targets (Ellner et al., 2012; Fencsik, Klieger and Horowitz 2007; Iordanescu et al. 2009), which may allow observers to use a more complex set of perceptual and cognitive processes to maintain selection of the target objects (Scholl & Flombaum, 2010). Motion

Author's personal copy

information might also be important for tracking without disruptions. For example, tracking performance is impaired by including motion information *within* an object (e.g., texture) that conflicts with the motion direction of the object itself (St Clair, Huff, & Seiffert, 2010), suggesting that this information interferes with an extrapolation process computed during "online" tracking. There are now multiple demonstrations that motion extrapolation plays at least some role in some types of tracking tasks. But the present results show that when observers track four target objects, there is no evidence that motion extrapolation helps bridge the disruption caused by object occlusion.

Acknowledgments We thank Todd Horowitz, Evan Palmer, Adriane Seiffert, and Dan Simons for helpful feedback and Andrea Cuddington, Pralle Kriengwatana, and Claudia Lau for their assistance in data collection. This work was supported by NSF CAREER Grant (S.F.) BCS-1056730.

References

Alvarez, G. A., Horowitz, T. S., Arsenio, H. C., DiMase, J. S., & Wolfe, J. M. (2005). Do multielement visual tracking and visual search draw continuously on the same visual attention resources? *Journal of Experimental Psychology: Human Perception and Performance, 31*(4), 643–667.

Awh, E., & Jonides, J. (2001). Overlapping mechanisms of attention and working memory. *Trends in Cognitive Sciences, 5,* 119–126.

Cheries, E. W., Mitroff, S. R., Wynn, K., & Scholl, B. J. (2008). Cohesion as a principle of object persistence in infancy. *Developmental Science, 11,* 427–432.

Comtois, R. (2004). VisionShell PPC [Software libraries]. Cambridge, MA: Author.

Ellner, S., Flombaum, J., & Scholl, B. J. (2012). *The necessity of extrapolation in multiple object tracking*. Manuscript submitted for publication.

Fehd, H., & Seiffert, A. E. (2008). Eye movements during multiple object tracking: Where do participants look? *Cognition, 108,* 201–209.

Fencsik, D., Klieger, S., & Horowitz, T. (2007). The role of location and motion information in the tracking and recovery of moving objects. *Perception & Psychophysics, 69,* 567–577.

Flombaum, J. I., & Scholl, B. J. (2006). A temporal same-object advantage in the tunnel effect: Facilitated change detection for persisting objects. *Journal of Experimental Psychology: Human Perception and Performance, 32,* 840–853.

Flombaum, J. I., Scholl, B. J., & Pylyshyn, Z. W. (2008). Attentional resources in tracking through occlusion: The high-beams effect. *Cognition, 107,* 904–931.

Franconeri, S. L., Alvarez, G. A., & Enns, J. T. (2007). How many locations can be selected at once? *Journal of Experimental Psychology: Human Perception and Performance, 33,* 1003–1012.

Franconeri, S. L., Bemis, D. K., & Alvarez, G. A. (2009). Number estimation relies on a set of segmented objects. *Cognition, 113,* 1–13.

Franconeri, S. L., Jonathan, S., & Scimeca, J. M. (2010). Tracking multiple objects is limited only by spatial interference, not speed, time, or capacity. *Psychological Science, 21,* 920–925.

Franconeri, S. L., Lin, J. Y., Pylyshyn, Z. W., Fisher, B., & Enns, J. T. (2008). Evidence against a speed limit in multiple object tracking. *Psychonomic Bulletin & Review, 15,* 802–808.

Franconeri, S. L., Scimeca, J. M., & Jonathan, S. (2012). *Maintaining selection of multiple moving objects*. Manuscript under review.

Gibson, J. J., Kaplan, G. A., Reynolds, H. N., & Wheeler, K. (1969). The change from visible to invisible: A study of optical transitions. *Perception & Psychophysics, 5,* 113–116.

Henderson, J. M. (1997). Transsaccadic memory and integration during real-world object perception. *Psychological Science, 8,* 51–55.

Horowitz, T., Birnkrant, R., Fencsik, D., Tran, L., & Wolfe, J. (2006). How do we track invisible objects? *Psychonomic Bulletin & Review, 13,* 516–523.

Intriligator, J., & Cavanagh, P. (2001). The spatial resolution of visual attention. *Cognitive Psychology, 43,* 171–216.

Iordanescu, L., Grabowecky, M., & Suzuki, S. (2009). Demand-based dynamic distribution of attention and monitoring of velocities during multiple-object tracking. *Journal of Vision, 9*(4, Art. 1), 1–12. doi:10.1167/9.4.1

Irwin, D. E. (1992). Memory for position and identity across eye movements. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 18,* 307–317.

Keane, B. P., Lu, H., & Kellman, P. J. (2007). Classification images reveal spatiotemporal contour interpolation. *Vision Research, 47,* 3460–3475.

Keane, B., & Pylyshyn, Z. (2006). Is motion extrapolation employed in multiple object tracking? Tracking as a low-level, non-predictive function. *Cognitive Psychology, 52,* 346–368.

Koch, C., & Ullman, S. (1985). Shifts in selective visual attention: Towards the underlying neural circuitry. *Human Neurobiology, 4,* 219–227.

Ma, W. J., & Huang, W. (2009). No capacity limit in attentional tracking: Evidence for probabilistic inference under a resource constraint. *Journal of Vision, 9*(11, Art. 3), 1–30. doi:10.1167/9.11.3

Palmer, E. M., Kellman, P. J., & Shipley, T. F. (2006). A theory of dynamic occluded and illusory object perception. *Journal of Experimental Psychology: General, 135,* 513–541.

Pylyshyn, Z. W. (2000). Situating vision in the world. *Trends in Cognitive Sciences, 4,* 197–207.

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

Scholl, B. J. (2007). Object persistence in philosophy and psychology. *Mind & Language, 22,* 563–591.

Scholl, B. J., & Feigenson, L. (2004). When out of sight is out of mind: Perceiving object persistence through occlusion vs. implosion [Abstract]. *Journal of Vision, 4*(8, Art. 26), 26a.

Scholl, B. J., & Flombaum, J. I. (2010). Object persistence. In B. Goldstein (Ed.), *Encyclopedia of Perception* (Vol. 2, pp. 653–657). Thousand Oaks, CA: Sage.

Scholl, B. J., & Nevarez, H. G. (2002, November). *Why so slow? The role of speed discontinuities in maintaining object persistence through occlusion.* Paper presented at the 43 rd Annual Meeting of the *Psychonomic Society*, Kansas City, MO. [Abstract published in *Abstracts of the Psychonomic Society, 7*, 19.]

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

Shioiri, S., & Cavanagh, P. (1989). Saccadic suppression of low-level motion. *Vision Research, 29,* 915–928.

St. Clair, R. L., Huff, M., & Seiffert, A. E. (2010). Conflicting motion information impairs multiple object tracking. *Journal of Vision, 10*(4, Art. 18), 1–13.

Author's personal copy

Volkmann, F. C., Riggs, L. A., & Moore, R. K. (1980). Eyeblinks and visual suppression. *Science, 207,* 900.

Vul, E., Frank, M. C., Alvarez, G. A., & Tenenbaum, J. B. (2009). Explaining human multiple object tracking as resource-constrained approximate inference in a dynamic probabilistic model. In Y. Bengio, D. Schuurmans, J. Lafferty, C. K. I. Williams, & A. Culotta (Eds.), *Advances in neural information processing systems 22* (pp. 1955–1963). La Jolla, CA: Neural Information Processing Systems Foundation.

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

Zelinsky, G. J., & Neider, M. B. (2008). An eye movement analysis of multiple object tracking in a realistic environment. *Visual Cognition, 16,* 553–566.